

*The Technique of Teaching*

THE APPLICATION OF  
PSYCHOLOGICAL  
TESTS  
IN SCHOOLS

N.E. Whilde

151.2  
WHI

B L A C K I E

1.20

Bureau Ednl. Res. Research	
DAVID	HING COLLEGE
Dated	30. 4. 58
Acces No	1400

151.2  
WHI



THE TECHNIQUE  
OF TEACHING

---

---

---

THE APPLICATION OF  
PSYCHOLOGICAL TESTS  
IN SCHOOLS

BLACKIE & SON LIMITED  
16/18 William IV Street, Charing Cross, LONDON, W.C.2  
17 Stanhope Street, GLASGOW

BLACKIE & SON (INDIA) LIMITED  
103/5 Fort Street, BOMBAY

BLACKIE & SON (CANADA) LIMITED  
TORONTO



# THE APPLICATION OF PSYCHOLOGICAL TESTS IN SCHOOLS

BY

NOEL E. WHILDE, M.Sc.(Psych.), A.B.Ps.S.

Psychologist in Charge,  
Sheffield Child Guidance Centre



BLACKIE & SON LIMITED  
LONDON AND GLASGOW

1955

151.2

WHI

NOTE: It should be clearly understood that the views expressed in this book are those of the author and in no way implicate the Authority for which he works.

# CONTENTS

## PREFACE

Page  
vii

CHAP.

### I. THE SCOPE OF PSYCHOLOGICAL TESTS 1

*The Value of "Rough" Knowledge. The Educationally Subnormal. The Academically Bright. Defects of the Old Type of Scholarship Examination. Tests in Schools Primarily Administrative Devices. The Value of Forming Categories. Reading Readiness. Testing the Individual Child. Intelligence and Attainment Tests. Matching Ability and Achievement. The Child in the Class. Class and School Surveys. Attainment Test Norms not Necessarily Binding. Tests and Educational Research. Prognosis. Diagnosis.*

### II. SOME BASIC CONSIDERATIONS 23

*The Difficulties of Subjective Judgment. Test Standardization. Age Factor. Brevity of a Test. Selection and Placing of Words on the Scale. Use of the Test. Standardized Reading Test. Timed Tests. "Headroom." Simplification of Measurement. Intelligence Tests and Attainment Tests. Narrow Scope of Intelligence Tests. Standardization of Intelligence Tests. Criticisms of Intelligence Tests. Validity and Reliability. Coaching. Intelligence Quotient. Constancy of Intelligence Quotient. Percentile Ranks.*

### III. TYPES OF TEST 46

*Intelligence Tests. Individual Tests. Pencil and Paper Tests. Verbal and Non-Verbal Tests. Some Practical Points. Other Pencil and Paper Tests. Performance Tests. Musical Ability. Attainment Tests. Interest Tests. Character and Personality Tests. Rating Scales.*

### IV. CHOOSING TESTS 68

*Priority for Intelligence or Attainment Tests. Transfer or Common Entrance Results. Reading Tests. Other*

*Language Tests. Spelling Tests. Arithmetic Tests. Intelligence Tests. Examiner's Booklet. Verbal or Non-Verbal Tests. Another Method of Selecting an Intelligence Test. Individual Tests. The Secondary Schools.*

V. DIAGNOSTIC TESTS 88

*Prior Need of Standardized Tests. Diagnostic Possibilities of Ordinary Standardized Tests. Specially Designed Diagnostic Tests.*

VI. DISTRIBUTION OF ABILITY 104

*Percentile Ranks. School Marks. Percentiles and Standard Deviation. Ideal and Actual Distributions. Distribution of Attainments. Rating Scales.*

VII. CLASS AND SCHOOL SURVEYS 128

*Individual Records. Class Records. Median and Quartiles. Age-Group Records. Age-Group Spread. School Surveys. Intelligence Quotient Tables. Attainment Norms.*

VIII. TEST MATERIAL 147

*Attainment Tests. Intelligence Tests. Special Mental Abilities Tests. Sensory-Motor Tests. Tests of Auditory Acuity. Tests of Visual Acuity. Tests of Colour Blindness. Interest and Personality Tests.*

BIBLIOGRAPHY 164

APPENDIX I 168

*The Educationally Subnormal.*

APPENDIX II 173

*Vocational Guidance.*

INDEX 177

## PREFACE

In this book, which is designed for teachers, I have tried to justify to some extent the use of psychological tests. The book, however, is not sufficient in itself to do this. I hope rather that the information it contains about tests will be adequate to guide the reader to a point where he will be able to use tests and feel that he is being repaid for the time and energy he has spent in using them. Tests suffer from those who claim too much for them as well as from those who are prejudiced against them. A clear conception of what tests can and cannot do is a necessity to-day. There are always people who take things up as a "stunt". This book is not intended for such people. Time and the collective judgment of teachers find out what is of real value in new educational ideas. I am not obsessed with the value of tests but I believe in them in the way that we all believe in that which "works".

There, are, of course, many excellent books already available on the subject of psychological tests, but most of them are rather technical and perhaps forbidding to the average teacher. In this book an attempt is made to see the subject from the teacher's point of view, and to give the essential facts in as simple a form as possible. I have also taken advantage of the opportunity to meet as fairly as possible the sort of criticisms which teachers make concerning tests.

Psychology is a new science and an even newer profession, but it is hoped that psychologists are overcoming the adolescent inferiority which has sometimes in the past led them to claim too much for their work and to be over-zealous in defending it from negative criticism. On the other hand, conservative principles cause many people to reject the psychologist and all his works out of hand. As one who has worked for a good



many years as a teacher before becoming a psychologist, I am fairly familiar with staff-room comments, often ribald, which refer to psychologists as half-crazy people who have the wildest ideas about education and are complete strangers to common sense. This book, I hope, will be a "go-between". I have tried to simplify the subject but not to talk down to teachers. I have listened to teachers and tried to see what their difficulties are. I shall feel rewarded if this effort on my part enables the reader to see that there is "something in it" after all.

Many teachers nowadays are asked to deal with test results. These are made available from the annual Transfer Examinations; Child Guidance reports are sent out; sometimes a teacher is called upon to administer tests for his Local Education Authority. All this demands some understanding of the principles and terminology of mental testing. As well as being concerned with this, however, I hope that teachers will find this book useful in developing their craft and increasing their efficiency.

There can, of course, be too much testing. The more discerning in this modern age see that the machine has mastered man in many ways. The ubiquitous omnibus can transform a tedious half-hour's walk into a leisurely journey of five minutes, but how many of us take advantage of the time thus saved to make life more dignified and spacious? On the contrary, we are often found running for the bus. On the following pages tests are presented primarily as administrative devices. We have already fallen into the habit of talking about the administrative *machine*. The teacher, however, should never allow tests and test procedures to become so important that they dominate him like a masterful machine. He should try a test and make it yield something. If he does not consider the result worth the time spent on it, he should give it up, but he must first be fair to the test. It is hoped that the following chapters will help him towards this end.

I am also conscious in this book of seeming to write as though I am interested only in formal subjects, especially the 3 R's. This, of course, is the result of writing a book on a

special topic. I have also referred to "the child" more often than I should have liked, but this has also seemed inevitable. A good deal of what I might call "training college psychology" suffers from the defect of over-generalization because of our modern mania for examinations and overcrowded syllabuses. The result of this is that students are not allowed enough time for individual study of individual children. I hope that the reader will not be content with my generalizations but will use the ideas here expressed so that he understands a little better, not "the child", but specific children.

It seems to me that there are three important factors in learning: ability; interest and instinctive drive; and technique or discipline. I use the word "discipline" in its wider sense of order, or compulsion to fulfil necessary conditions. Thus, in arithmetic the discipline of the subject demands that we should first learn basic number combinations, that we should be able to subtract before we can divide, that we should understand the principle of L.C.M. before we can make any progress with fractions. A good teacher has the technique of all this clear in his mind and introduces the right step at the right time. The bad effect on a child who misses some vital step through absence, especially in the early stages, is well known. In the past the main criteria were "what to do" and "how to do it"—the syllabus and the technique. So it appears that for a long time the third of my important factors was often the only one considered.

As far as interest was concerned, children were too often regarded as passive quantities, vessels to be filled up, and this process of "filling up" often degenerated to one of "knocking in". As for the instinctive drive, this all too often took the shape of fear—the child's fear of the teacher, the teacher's fear of the inspector. We like to think that those days are gone. We now understand the importance of the child's nature. We use activity methods, "play ways", competitive teams, and co-operative groups.

I placed ability first among my three factors. Yet it has always been—and still is—the least considered. In our schools to-day many bright children are bored, many dull



children are confused, and many teachers wear themselves out and become dispirited because they have not a proper understanding of the actual *ability* of their children. One of the aims of this book is to bring this first factor of ability into greater prominence in the teacher's mind. I place it first, not because I think it is more important than interest and drive or than a sound technique, but because it naturally comes first. The second and third factors may be misapplied without adequate and fairly accurate appreciation of the first.

It is very necessary to mention here that there is still controversy and disagreement about theoretical matters relating to intelligence and special abilities or factors, but this does not concern the average teacher. I have tried to take up the point of view of common sense. We do not refuse to use soap because we do not understand why it is so efficient in removing dirt; nor would the position be very much affected if we thought it worked for quite a wrong reason. We should still use it and be glad to have it. With proper safeguards and within limits mental tests work; that is, they produce results which can be observed, studied, and used to influence our behaviour advantageously. That there will be better tests in the future which will come from an increased theoretical knowledge is probably true, but the time now seems ripe for a considerable extension of mental measurement by teachers, based on such findings of psychologists as are generally agreed on at the present time.

Finally, this is not a book *of* tests but a book *about* tests. It is hoped that when his Local Education Authority requires him to give tests—and it is increasingly the practice of Authorities to ask teachers to do so—the reader will tackle the problem equipped with a knowledge that will enable him to carry out his task competently and with a knowledge of what he is doing and why it should be done. He should be able, too, to appreciate the difference between intelligence tests and attainment or other scientifically constructed tests, to read and interpret a psychologist's report, and to use with adequate understanding that part of the modern Scholar's Record Card which involves test results. Over and above this, however,

it is hoped that many teachers will find that psychological tests can help them to understand their own problems and adapt themselves intelligently to their tasks. Perhaps the point of view which they will find most useful of all is that expressed in Chapter I, namely, that tests are fundamentally administrative devices.

The reader will notice some repetition here and there. This has been deliberate, as it helps to make each topic intelligible by itself. It is thought, too, that the reader coming to the subject for the first time will welcome what are in effect revisions of points made earlier in the book or in other contexts.

N. E. W.

## CHAPTER I

### THE SCOPE OF PSYCHOLOGICAL TESTS

There are two attitudes which "lay" persons are apt to adopt towards "specialists". At their extremes one is an attitude of reverence and deification and the other is one of mistrust and almost rebellion. These two attitudes are observable towards psychologists and they are also found in relation to that branch of the psychologist's work which forms the subject matter of this book, namely, psychological tests. It might be said that these two extreme attitudes are *both* a mark of intolerance, for neither type makes provision for uncertainty, for questioning, or for what might be called approximate knowledge. Reflection shows, however, that man's ascendancy over his environment has proceeded by steps which have involved the use of partial truths, of approximate knowledge, though there have always been the over-enthusiastic (and perhaps gullible) and the scorners at every stage. As regards the present topic, the author hopes to find his reader tolerant of the admittedly imperfect state of development which psychological testing has reached.

#### **The Value of "Rough" Knowledge.**

We should always remember that we can never establish absolute certainties by means of psychological tests. In this book it is hoped to face up to some of the uncertainties, and yet to show that on the whole the tests are very valuable. We must avoid expecting easy certainties but proceed by practical steps to make as certain as we can. We must also avoid thinking of psychological terms as magic words. Many people use the



terms "I.Q." (intelligence quotient)<sup>1</sup> and "M.A." (mental age) in a loose indefinite way probably because they do not fully understand their meanings themselves. For instance, one can say *A*'s I.Q. is 100, but it is much better to say Mr. So-and-So obtained an I.Q. of 100 for *A* on a certain test. It is then possible to form an opinion as to the value of the result, if one knows the test and Mr. So-and-So. It is also equally important to know how old *A* was at the time of testing, or how long ago it is since the test was administered. Again, if a test result is different from our native judgment, it is always wise to give a second test or to obtain a second opinion about the child. It is my experience that in the vast majority of cases the test is sounder than our own judgment, but although we may believe this we should always be true to what we think is right until we are proved wrong.

Assuming we have now obtained as good a measure of a child's intelligence (expressed by his I.Q.) as is humanly possible, what do we know about him? We know nothing that will throw any light at all on his temperament or character, nothing that will tell us whether he is likely to be good at games or good with his hands, but we do know what his chances of being successful as a "scholar" are likely to be. Here again we know this only roughly, but rough knowledge can be useful.

### The Educationally Subnormal.

It is useful to know that certain children are likely to be unable to profit by the ordinary school provision made for them. In other words, if certain children are relatively so unintelligent that they cannot be adequately taught under ordinary conditions, and they need special education in a special school, it is useful to be able to pick out these children by means of tests. There are, of course, some difficulties. Our knowledge is not so exact that we can always be sure of

<sup>1</sup> Readers approaching the subject of psychological testing for the first time might well first read Chapter II, page 38, for definition of I.Q.

accurate selection, especially at the border-line, but here to supplement the test we can use information on other important factors, such as emotional stability and school record. By means of psychological tests it is possible, with a high degree of accuracy, to select these educationally subnormal children at the age of 7 or 8 years and give them special educational treatment.

### **The Academically Bright.**

In the same way children with sufficient aptitude for a grammar school education can be selected in a very satisfactory manner by means of psychological tests, usually by a combination of intelligence and educational tests. Again there is most difficulty in dealing with border-line cases, but here, too, additional refinements can be applied to these cases so that the most suitable of almost equally intelligent children can be transferred.

### **Defects of the Old Type of Scholarship Examination.**

The method of using psychological techniques in the Common Entrance or Transfer Examinations is much superior to the old examination method. A study of the results of an "old style" arithmetic scholarship paper showed that about half of the problems set were done correctly by all the candidates likely to be selected and were therefore of no value whatsoever except as "warming up" exercises. Of the other half of the problems, several were "catches" and depended a good deal on chance and coaching rather than real arithmetical skill and intelligence. All these problems involved long calculations in which a slip in the early stages of the working meant that no credit was obtained.

The whole system was thoroughly unsatisfactory and was successful only because there were at that time so few grammar school places available. Competition was so keen that, once the poorest candidates had been eliminated, a pin could quite easily have been the means of selecting children who would do



fairly well at a grammar school from the remaining available children of high ability. There are now in general many more grammar school places available; in some areas from 15 to 20% of the secondary school population go to grammar schools. Such a system as that mentioned above would almost certainly break down under these conditions. The "raw" arithmetic scores in the above-mentioned test were added to the "raw" English scores to obtain a final order, although there were many more marks given for English than arithmetic. Thus a child relatively good at arithmetic and not so good at English was penalized in the examination. As is usual in such examinations an allowance was made for age, a certain number of marks being given for each month the child was younger than the maximum age. Here again the allowance was entirely arbitrary—decided by a method of guesswork which may have favoured the younger children or may have penalized them. Now with the modern statistical procedures which the educational psychologist uses, together with his standardized<sup>1</sup> attainment and intelligence tests, all the more serious of the objections to the old type of test have been overcome. This is a most important development now that it is a statutory duty for each Local Education Authority to educate each child according to his age, aptitude, and ability, for now *every* child must be examined at the appropriate time and directed to a suitable secondary school. Formerly the Scholarship Examination catered only for those children whose parents were interested and for whom the Local Authority thought fit to provide grammar schools.

There is still great variety in the selection methods of the various Local Authorities and a great deal of experimenting is taking place. Attempts are being made to take into account the child's educational record and to give weight to character factors. Criticism can be levelled at the procedures so far devised, but the point is that, imperfect as our methods are, they are not bad methods. Some injustices there are, but it must be remembered that, however good any system of transfer

<sup>1</sup> See Chapter II for meaning of "standardization".

is, there will always be border-line cases where a child is fit for either this type of school or that one, and the final choice must be to some extent arbitrary. By a review of the child's progress during the first or second year at the secondary stage errors of selection can to some extent be rectified. Mention of Technical School selection has been omitted deliberately. This topic will be of increasing importance in the future, but as there is still controversy about the tests to be used and the age at which they are effective, and as their consideration does not affect the present argument, they have been passed over. Reference to such tests will be found in Chapter III.

### Tests in Schools Primarily Administrative Devices.

These two examples, which refer to the directing of the very dull and the very bright to appropriate schools, have been deliberately chosen here to call attention to the point of view implicit throughout this book that psychological tests are primarily administrative devices so far as schools are concerned. Their function is to sort out children into categories. Upon reading this statement many teachers will no doubt feel that their worst fears are now confirmed. They may argue as follows: "Children are to be labelled! We have already too much mechanization and direction. We grant that it may be all right to have a large-scale examination for selecting an appropriate secondary school, and that educationally sub-normal children should have special treatment, but can't it stop there? Why try to bring the process into the school itself?"

My answer is that we already label children. We have an "infant", a "junior", and a "secondary" scholar. Within each school, too, we have our forms and classes. No one would suggest putting a 5-year-old in the sixth form at a grammar school. It is all a matter of convenience and common sense. Even within a single class the discerning teacher often arranges the children in groups for different subjects in order to achieve the maximum results from the whole class. This is not thought



of as "labelling", largely because such arrangement is not static. A child may be moved from the bottom to the top "set" if he improves, and the reverse procedure may be adopted if he does not. We need beware of the "labelling" process only if we adopt the rather immature attitude previously mentioned of giving words too much "magic" and of assigning to an I.Q. an importance and meaning far beyond what is due to it. A perusal of Chapter II should help to give perspective to this matter.

### **The Value of Forming Categories.**

There are many reasons why sorting out or categorizing is both desirable and economical. For example, it is obvious that a policeman must have a reasonably good physique and height above the average. Whatever other qualities are desirable, these two are essential, so that minimal standards can be set up and a category formed. It would be fatuous to examine a candidate for the police force, require him to give proof of good character, level of scholarship, etc., and then finally reject him because he was only 5 feet tall. It would be even more foolish to appoint him and then dismiss him because his height made him unsuitable. Yet this is just the sort of thing that goes on in schools all the time, and as a result of our refusal to make categories and apply minimal standards we constantly find a fantastic waste of energy and an equally fantastic amount of frustration.

In the past no one was to blame for this because no one knew what minimal standards could be applied. It is the use of psychological tests which has helped to provide these.

### **Reading Readiness.**

Let us take for example the question of reading readiness. It can be shown that a certain amount of intelligence is necessary before children can learn to read with comprehension. (I must emphasize that I am speaking about *reading comprehension*, for infant teachers might be ready to contradict me if they



think only of *word recognition*. Whilst word recognition is an essential part of reading, in itself it is only what has been described as "barking at print". Many teachers or parents will have had the experience of knowing a child who is so familiar with a primer that he can read all the words in it, but who is lost if given another book to read. This is not reading; it is a habit or trick.) The lower limit for beginning reading is a *mental* age of approximately 6 years.<sup>1</sup> If we examined an average group of children,<sup>2</sup> that is, a group composed of children of the usual varied standards of intelligence, we should find that only half of them reach a mental age of 6 by the time they are chronologically 6 years old. The brighter children reach this critical level earlier, but almost a quarter of the children have not reached a mental age of 6 by the time they are  $6\frac{3}{4}$  years old. The dullest 10% are over  $7\frac{1}{2}$  years old before their mental age is 6.

What useless slaving goes on in many schools with this dullest quarter! Many children become so discouraged in these early years that they become almost inoculated against learning to read by their early, constant, and continued failure. The application of an intelligence test and the setting up of categories—"Ready to read", "Not ready to read"—would yield a handsome educational dividend on the time and trouble spent on testing in the Infant School. It has already been emphasized (and this point will be dealt with again in other parts of the book) that there should be no hair-splitting about these categories. It would be foolish to say: "This child cannot read to-day; to-morrow he must be reading." We are applying only approximate knowledge. Before reading proper is begun all children need preparation for reading,<sup>3</sup> and whenever a child appears ready to read he should be given the chance to do so. The main thing is that we should establish

<sup>1</sup> American investigators have fixed this mental age at  $6\frac{1}{2}$  years. A well-documented summary of the research done and a critical appraisal is given in *Studies in Reading*, Vol. I. See especially pp. 69-86.

<sup>2</sup> Normal distribution is discussed in Chapter VI.

<sup>3</sup> See *The Psychology and Teaching of Reading* (Schonell).

a good guide to what can be expected of children in general, and in particular the duller 10 to 20% which so often form the "hard core" of difficulty for the Junior and Secondary Modern School.

Up to this point we have had in mind an "average" school and its problems in relation to reading readiness, but of course all schools are not average, and many schools have a higher proportion of dull children than has been indicated so far in this discussion. Some infant departments might well decide that their function should be largely to elaborate and continue the work of the Nursery School or class, and leave most of the formal 3-R work to the Junior Department. This policy may seem surprising, but where a school is fed from a socially very poor area it might be necessary for a large proportion of the children in it. The junior department fed by such a school might resent it, but if a competently conducted testing programme showed that about half the children in a particular school do not reach a mental age of 6 until they are 7+, then, so far as these children are concerned, formal 3-R work should be postponed until the age of 7+. This method would have three definite advantages: the children would understand what they were doing and would find the work more enjoyable, the teacher would have the satisfaction of making real progress, and there would be fewer "reading defectives" at later stages.

It is clear that the plea for a postponement of reading for the duller children in *some* schools is reasonable when it is shown that in certain schools the mean I.Q.'s are as low as 90 and even 86. Thus, since  $I.Q. = \frac{M.A. \times 100}{C.A.}$ , if we stipulate a minimum M.A. of 6 for reading and take an I.Q. of 90 we have  $90 = \frac{6 \times 100}{C.A.}$  which yields  $C.A. = 6$  years 8 months; with an I.Q. of 86, in the same way,  $C.A. = 7$  years. But half the children will have chronological ages greater than this. The border-line figure at which we reach the "educationally subnormal" (who should not be in an ordinary



school) is an I.Q. of 70—in other words, the dullest children with which an ordinary school can be expected to deal have I.Q.'s of about 70. These children do not reach a mental age of 6 until they are over  $8\frac{1}{2}$  years.

### Testing the Individual Child: an Administrative Matter.

At this point it might be well to assert that, although so far I have discussed the use of intelligence tests to distinguish groups of children, the principle still holds good in the case of individual children. It will be seen from Chapters II and VI that the standardization of a test involves placing a completely representative group of children on a graduated scale, so that the administration of an intelligence test to *one* child and the calculation of his mental age or intelligence quotient, enables us at once to say where on that scale he should be found. We can thus put him in a particular category, and this ease in classification is a result of the work which psychologists have done in devising and standardizing the test. Moreover, when other research has established a *maturation level*<sup>1</sup> (which in the case of reading means the mental age at which a child is *ready* to read) our categories will become significant instead of entirely arbitrary, and will be just as useful for one child as for a group.

<sup>1</sup> All educators should appreciate the principle of "maturation". This refers to the maturing of certain innate abilities *without* training or experience. Experiments with identical twins have helped to provide evidence for this. One receives a more or less lengthy period of training, for example in climbing, while the other remains untrained, and as far as possible deprived of the opportunity of learning. It is then found that, at a given age, the twins can do the task almost equally well, showing that the ability matured of its own accord, or was a product of the maturing personality and not a result of the training. Similar experiments with babies and young children (with control groups where necessary) show that the ability to fasten buttons or use scissors, for example, is as much a product of *maturation* as training. A more exact knowledge of maturation levels would probably result in the introduction of various topics into the syllabus. It might be necessary to postpone some, and to remove others altogether, as the research of Macauley (see page 18) into the question of formal grammar in the Junior Department would seem to indicate. *From Birth to Maturity* by Charlotte Bühler deals extensively with the factor of maturation. It is also discussed in *Personality* by G. W. Allport.

### Intelligence and Attainment Tests.

Reference has already been made to *intelligence tests*<sup>1</sup> which serve to indicate a *mental age*. *Attainment tests* give attainment ages, e.g. reading age, spelling age, arithmetic age; and just as the *mental age* of a child is a measure of the intelligence standard which he has reached, the *attainment age* is a measure of the standard of reading or other ability which he has reached. There is a fundamental difference, however, between intelligence and attainment, in that intelligence is an inborn capacity and develops at its own intrinsic rate, while attainment is an acquired capacity. Psychologists have always striven to make intelligence tests as independent of the influences of environment as possible, or to use those environmental factors which are common to most children in such a way that they do not become disturbing factors. It is obvious, however, when we consider attainments, that these are influenced very much by environment: the amount of time spent on a subject, the quality of the teaching, the size of the classes, as well as other factors, must have some effect on what is achieved in any subject. As will be explained in Chapter II, attainment tests endeavour to minimize extreme environmental influences, and to give results which can in fact be achieved by children chosen from representative types of schools.

### Matching Ability and Achievement.

With the help of efficiently standardized intelligence and attainment tests we may now try to compare attainment with capacity, whether it be in the case of the individual child, the class, the age-group, or the school. Ordinarily it would seem reasonable to expect them to correspond. Let us for the moment consider the individual child. If he is an average child and we find his attainments are those usual in a child of his age, we may consider that all is well. At 10 years old, because he is an average child, we should expect his mental

<sup>1</sup> Types of test are discussed in Chapter III.



age to be 10 years, and his attainment ages to be about the same. If he is dull, perhaps two years retarded mentally, his mental age will be 8 years, and if we find his attainment ages to be about 8 years, we may again conclude that all is well. Similarly, a bright child will have a mental age higher than his own age, and we may expect attainment ages to correspond.

In actual fact, however, attainments rarely match intelligence with precision, and to expect them to do so would suggest that the educational process is merely a machine, for such an expectation would overlook the fact that, while intelligence is the predominant factor in learning, it is not the only one. In reading, for instance, the child's cultural background is important, as is also his capacity to discriminate word patterns, both aurally and visually, for these two capacities appear to be specialized mental abilities, beyond auditory and visual acuity, that is, over and above what we should ordinarily regard as seeing or hearing defects.

If therefore a child's attainment age is approximately equal to his mental age, he is progressing satisfactorily. If there are considerable deficiencies, further investigation may be necessary to discover the specific cause of the backwardness, and some of the diagnostic tests mentioned in Chapter V may be required. It is not always wise, however, to assume that a measure of retardation has some complicated cause. It may have, but it is wise to eliminate the simpler factors first. Sometimes bright children are retarded in a particular subject, merely through lack of educational opportunity. Consider, for instance, a bright child of 7 years with an I.Q. of 130. Such a child should have a mental age of  $7 \times \frac{130}{100} = 9.1$  years. If

his attainment age in spelling is only about 7 years, he is two years retarded compared with his mental age, but since he has not been doing formal school work for more than two years at the most he has perhaps not had the opportunity to practise spelling. The same situation will frequently be found with the reading ability of a bright child whose home does not



supply a "bookish" background. Although he may have the desire, the child is deprived of the opportunity to do the extra reading which would help his reading attainment age to catch up with his mental age.

The position would be entirely different in the case of a 10-year-old child of average intelligence whose attainments had reached only those of an average child of 7 or 8 years. A comprehensive investigation would have to be made before the matter could be adjusted, but here again there is the possibility that his difficulty may have been caused by serious absences at some critical time.

For example, a certain grammar school boy of very high intelligence had been given up as hopeless by his mathematics master. Some investigation revealed that, never having been taught basic addition and subtraction, he relied on guesswork. It may seem hardly credible, but he had in fact been doing this since he was about 7 years old and had been unable to make any headway in the subject. He had such a strong feeling of inferiority that he could not of his own volition find a way out of his difficulty. He had been punished and held up to ridicule so often that he really felt he must be the fool he had so often been told he was. Naturally in algebra he was even more confused, for if he could not make  $6 + 7 = 13$  he could not make  $6x + 7x = 13x$ . When it was explained to this boy what his trouble was, and when he was told that there were only 100 addition combinations, including the noughts, up to 9 and 9, he was very anxious to put the matter right. Aided by his teacher and his own determination, he mastered all these and the corresponding subtraction combinations in two or three months, and could then tackle and overcome his other difficulties in mathematics.

### **The Child in the Class. Teachers' Difficulties.**

There is one further point concerning individual backwardness which is worth mentioning. This is the question of the "class level". It is necessary to emphasize again that tests

should be used with discretion and common sense. It is clearly impossible to give each child individual tuition, and every teacher of a large class must be depressed and disheartened by the realization of how few minutes can be given to any one child for individual attention in one lesson or even in a whole week. It is always more difficult to deal with the child who is outside the middle range of the work of the class, and I should be doing a disservice to teachers if I let it be assumed that the practical difficulties of the classroom are not appreciated, or that every child can be easily taught up to his capacity. This is only an ideal towards which one must strive. Nevertheless, the objective information obtained from tests does enable the teacher to see just what the child's problem is, and can even at times be a protection for the teacher, who can show his tests results to those in authority in order to demonstrate the nature of the problems with which he is expected to cope.

I recall examining several children in a class and finding most of them to be mentally defective, or, as we now much more appropriately say, "educationally subnormal", for they were not really educable in the ordinary school. I discussed the problem of these children with the class teacher, who told me that she had a very dull class indeed. As this was her first post she had no background of experience to guide her in the tackling of her difficulties. She was depressed by the problems which faced her, and felt that as a teacher she had failed. There had been no real recognition of her difficulties, for the syllabus she was expected to cover was obviously based on the capacities of average children. No young teacher, fresh from college, should have been given an assignment which would have been a hard task even for an experienced teacher. In one's enthusiasm to demonstrate what can be done for children under ideal conditions it is easy to be guilty of failing to appreciate the concrete situation. There is a wide gap between what can be done under average existing conditions and what might be done with small classes in favourable circumstances. I therefore maintain, in considering the individual child and our attempts



to match attainment to capacity, that it is always important to relate the problem to the background, and what I have called "class level".

I have suggested in this book that we should not expect too much of tests. I should like also to suggest that we should not expect too much of ourselves as teachers. I know from experience that those who begin by expecting too much of tests finally give them up. On the other hand, having used tests to bring to light what *should* be done for children, we should keep our feet on the ground, and not imagine that we can achieve everything at once. We see that it is necessary to keep a balance between the needs of the child and the teacher's capacity to stand "wear and tear". But though psychological testing can set many problems in startling relief, these do not represent the sum total of all teaching difficulties. The young teacher already mentioned was not in a state of despair because of any testing which had been done. It was later that testing showed that her problems were partly those of bad organization and of failure on the part of others to understand the magnitude of her difficulties. She had been taking on herself, or had been given, responsibility for an almost impossible task. Tests can be used to protect the teacher from injustices of this kind.

### **Class and School Surveys. Policy Making.**

It has been pointed out that there is a place for attempting to match achievement to ability not only in the case of the individual, but also in the case of the group. This can be achieved by class or school surveys, which are dealt with in detail in Chapter VII. Such surveys are also very helpful in policy making. Consider, for instance, the question of "free activity" methods versus the more mechanical aspects of schooling. Miss D. M. Gardner in her valuable book *Testing Results in the Infant School* has amply justified the policy of using informal methods at the infant level. At the junior stage there is a wide field for the application of tests to discover the effects of "policy".

It is difficult to date precisely "movements" in education, but early in the present century there emerged a healthy tendency to make arithmetic schemes more realistic and practical. This unfortunately led some teachers to neglect the more mechanical aspects of the subject, and one sometimes felt that the "baby had been thrown out with the bath water". This criticism is still valid to some extent, for the administering of group tests of mechanical arithmetic frequently reveals many children at the secondary stage busily manipulating their number combinations by finger counting. There are now textbooks such as the *Right from the Start* and the *Beacon* which give adequate recognition to the fact that the only solid foundation for learning arithmetic is an almost automatic knowledge of number combinations. There is no doubt that these textbooks must be used intelligently, and the teacher may feel uncertain about keeping an adequate balance between the "mechanical" and "problem" or "practical" aspects of the subject. This can be settled by using standardized mechanical and problem arithmetic tests, and comparing the results.

Where this is done quite surprising results are occasionally obtained. In one primary school a difference of over two years was found between "mechanical" and "problem" results, much to the surprise of the head teacher. This situation had arisen from a keenness to present the subject in a certain way, a policy which had caused quite a serious neglect of the mechanical side of the work.

In reading, again, teachers are sometimes prone to put a great deal of faith in certain methods, or a certain series of readers. A number of standardized tests, such as word recognition, comprehension, speed, or accuracy, would provide an objective measure of the success of the favoured method, and the sort of balance being achieved among these various aspects of reading.

There are other ways in which the results of policy can be tested. In some schools spelling is taught only incidentally, in others it is given a place in the time-table and is dealt with



systematically. Taking into account the erratic nature of English spelling, it is my view that for most schools the systematic way is advisable. Opinions, however, do not always carry us very far. The application of a well-standardized spelling test would reveal at once whether the system in use was at least adequate, even though it might not show whether it was an ideal one or not. The head teacher might also obtain some guidance in solving the vexed question of how much time to allow for each of the various subjects in the time-table. If, for example, the standard of arithmetic in a particular school is well above the average for the general level of ability in the school, it might be possible to allow extra time to another subject at the expense of arithmetic.

### Attainment Test Norms not Necessarily Binding.

Finally, before leaving the question of tests and policy, I would suggest that teachers should always bear in mind the fact that attainment test norms<sup>1</sup> are derived from children's actual work. It is true that the children are chosen so as to be as nearly as possible *representative*, but apart from their natural limitations what is in fact achieved is to some extent due to the syllabus and their teacher's expectations. The teacher, in this respect, should be free to express his opinion as to whether the test is above or below the capacities of his pupils. (We are for the moment speaking of average children.) Improved teaching techniques may make it possible for children in the future actually to achieve more. The important thing is that, given a standardized test, we can compare our children with a fair sample of other children. What action is taken thereafter should be at the teacher's discretion. It should always be remembered that the test should be the teacher's servant and not his master.

<sup>1</sup> A certain score on a test can be turned into an attainment age. The table of scores for a test, and the corresponding attainment ages, are referred to as the Table of Norms for that test.



## Tests and Educational Research.

The scope of testing cannot be dealt with exhaustively in a book of this type, but it would be a pity if in attempting to cater essentially for the average teacher no hint was given of the value of tests for the teacher interested in the more modest aspects of educational research.

For example, a head teacher, who taught arithmetic to a group of backward boys, found that most of the group (in "Junior II") could not understand the process of "equal addition" in subtraction, where what is usually called "borrowing" is involved. Some months later the process was taught again, and it was found that some of the boys were able to grasp it. This raised the interesting question for the teacher, whether a "maturation level" is not involved. Is there a relation between mental age and ability to grasp this process? The application of an intelligence test, and a suitable attainment or diagnostic arithmetic test, might reveal that it would be more profitable, or even essential, to introduce this process in "Junior III" in that particular school.

There must be many ways in which observant and enterprising teachers could apply themselves to such minor schemes of research. For instance, there is the amusing example, reported, I believe, by Ballard, of the respective values of sleep and coaching in improving the arithmetic of backward children. As it was objected that the children in a certain school always seemed tired, a controlled experiment was devised in which one group had extra coaching in arithmetic, and the other had equivalent periods of sleep. At the end of the experiment the work of the children who had had the extra sleep showed as great an improvement as that of the children who had had the extra tuition.

Watts<sup>1</sup> has recently claimed that a good deal of history teaching is misplaced. He asked a large number of children what questions they would like to ask about a man, whose

<sup>1</sup> *The Language and Mental Development of Children.*

statue they saw, and of whose existence they had never heard. At 7 years old they were as much concerned about the material of the statue as about anything else. It was only with the increasing age of the children that they began to ask about the reasons for the statue's existence, and the man's claim to fame. The low score on the same author's English History Test<sup>1</sup> also gives support to his contention. Readers interested in research into teaching the humanities will find many hints in Watts's pioneer volume.

A more ambitious piece of research by W. J. Macauley<sup>2</sup> shows very conclusively that formal grammar should not be taught in the Primary School, as it is merely time wasted. So far as Secondary Schools are concerned, there is strong evidence of the advisability of teaching it only to the brighter children who are taking a purely academic course. The results of this research should also cause teachers of foreign languages to question the wisdom of the *grammatical* approach to their subject, especially in the early years of the secondary course.

The idea that backward children are necessarily good at practical work has now been proved erroneous. That they need a more practical or concrete approach is no doubt true, but too much can be claimed for the results of practical work with backward children. In my opinion this movement towards more practical work for dull children has become caught up in the great extension of craft work proper, during the past twenty years. Thus we find many "C" and "D" stream children struggling with technical drawing in Secondary (Modern) Schools because it is rightly felt to be necessary to the understanding of such crafts as woodwork and metalwork. In my opinion the necessary abstraction involved in technical drawing is beyond the capacity of the duller children. I also wonder whether many children of 11+ have the necessary physical capacity for much of the tool control necessary in these subjects. These problems could be the subject of a very useful

<sup>1</sup> *ibid.*    <sup>2</sup> *British Journal of Educational Psychology*, Vol. XVII, Part III.



piece of research. The investigator would have to measure the strength of grip, steadiness of hand, etc., and such work would lead him far beyond the scope of this book. This type of problem, however, is in urgent need of solution, and I am sure any teacher working for an Education Authority which employs a psychologist would receive both advice and encouragement from the expert.

### Prognosis.

It has been implicitly assumed in earlier remarks that it is possible to prognosticate success in the formal aspects of school work by means of intelligence tests. It is fairly easy for a psychologist to estimate whether a child of 3 or 4 years old is likely to be dull or bright, though one would hesitate to be too definite at so early an age. Intelligence tests are often very useful in individual cases. For example, in the adoption or boarding-out of a child a knowledge of his potential ability would be most helpful in deciding on a suitable home for him. Similarly, new children arriving at a school from another area can be tested and correctly placed. Knowledge of a child's position in class at another school is not always adequate as the former school may have quite a different educational standard from the new one. Serious disturbances in a child's life have been known to follow a change of school, where incorrect judgment of capacity has been made.

There is, for example, the case of a 9-year-old boy who, owing to a physical disability, had never been to school. His parents then desiring to send him to school asked me to give an opinion as to his mental ability and to teach him the rudiments of the 3 R's if I found him to be mentally fit. I found his I.Q. to be in the lower 70's, and his mental age to be 7+. While his I.Q. was only just within the educationally normal category, he was in fact very dull. His mental age, however, was such that I felt that he could make some progress, and I therefore began to teach him. He was very keen to learn, with the result that within the space of six hours he mastered a



reading course designed to cover a period of two years. These six hours were, however, the sum of a number of approximately twenty-minute periods spread over several weeks. Of course he did not continue to progress at this rate as he soon reached an attainment age matching his mental age.

All examinations for transfer to suitable secondary schools are conducted as a means of prognosis, for only such children as are likely to succeed in the type of school to which they are assigned should be so transferred. There are now, however, tests available which will forecast a pupil's chance of success in certain subjects. Earle's *Tests of Ability for Secondary School Courses*<sup>1</sup> are suitable for grammar school children of 11-12 years, and, according to the score obtained, these tests show the child's chances of passing, later on, certain subjects in the School Certificate.<sup>2</sup> For example, for a score of 60 and over on the "Ability for Algebra" test there is a 100 to 1 chance in favour of the child passing in the School Certificate, whereas for a score of under 36 the chances are 20 to 1 against such success. The probability of pass or fail for various intermediate scores is also given. There are, too, tests of ability in science, geometry, and English. The science test gives probabilities of pass or fail in School Certificate chemistry, physics, or biology.

### Diagnosis.

It will be realized that these tests of ability, since they also differentiate between abilities, are diagnostic, too. Earle has in addition prepared his "Duplex" series of ability tests which yield estimates of both a child's general ability (I.Q.) and of differences which may exist between one type of mental process and another. Thus the Duplex No. 2 test gives not only an I.Q. and mental age, but also distinguishes between the

<sup>1</sup> University of London Press. See also *Reconstruction in the Secondary School*.

<sup>2</sup> The probabilities will not necessarily be applicable to the current "General Certificate of Education", but they will still distinguish between pupils and form a good guide as to their prospects. The effect of the changed examination arrangements on the forecasting ability of the tests can be gauged if teachers will keep records of children's scores and subsequent success.

efficiency of the mental abilities involved in "Vocabulary and Use of Words", "Language Study", "Science in Everyday Life", "Mathematics", and "Effective Approach to a Practical Problem". These tests will obviously be of considerable use both in discovering why children do not do well in certain subjects, and in directing them to more appropriate courses.

In diagnostic work general ability is the first factor to be considered, followed by an examination of special abilities, but, as already indicated in the case of the boy who failed in mathematics,<sup>1</sup> our knowledge has to extend to disabilities which may be specific and arise from some lack or defect in teaching, or from weakness, not in ability, but in the perceptual processes. The tests which have been developed to analyse specific difficulties and disabilities are of very great value. It is true that diagnostic tests can take up a great deal of time, but teachers will lose nothing by being familiar with them. Chapter V contains detailed information about diagnostic tests. Almost every school, including Grammar Schools, would find this test material extremely useful, for the possession of high academic ability is no guarantee of automatic freedom from specific disabilities, and many grammar school children in the lower forms could be helped by a careful analysis of their weaknesses.

Many schools would benefit by the appointment of a specialist teacher, preferably with a university diploma,<sup>2</sup> who has a sound knowledge of diagnostic tests. The idea of a special class for backward children is not new and is often a valuable part of a school's organization. What is less frequently found is an adjustment class for *retarded* children. Such a class caters, not for the backward child who is backward solely because he is intellectually dull, but for the child who for some reason or another is not progressing according to his capacity. He

<sup>1</sup> Page 12.

<sup>2</sup> Local Education Authorities are now empowered to pay a teacher's salary for a year while he takes such a diploma.

could attend one of these classes for just the length of time necessary to remedy his particular disability, and could well be dealt with by such a specialist teacher as is mentioned above. These classes can work reciprocally with Child Guidance Centres to great mutual advantage.





## CHAPTER II

### SOME BASIC CONSIDERATIONS

#### The Difficulties of Subjective Judgment.

Some of the values of mental measurement have been discussed in Chapter I, and my opinion that it is best to regard tests as administrative devices is dealt with there. In this section of the book the value of testing or measuring, with the special provisos dealt with in the previous chapter, is taken for granted. If then we believe that measuring is worth while we must ask ourselves how it can be achieved. The first problem is obviously to find a unit or yard-stick. Teachers with a great deal of experience will often claim that they can size up their pupils, both as regards ability and achievement, as well as, or better than, any test devised by a so-called expert. Now this is often true, so long as the teacher's judgments are related to the particular environment where he has gained his experience. There are, however, very considerable differences between schools (see Chapter VI), and when it is necessary to compare children from different schools, as in the Common Entrance or Transfer Examinations of to-day, these subjective judgments have a very considerable environmental bias which would result in a great deal of injustice if given too much weight. I have known teachers claim far too much for certain children on account of their intellectual superiority, when this superiority has been related only to the other children in a particular school, which has been in fact mediocre or poor so far as the innate intellectual endowment of the children is concerned.

An illustration of the limitation of subjective judgments was provided by a school in a "good" social area which had

rapidly deteriorated. The children in this school came from well-to-do homes and were of superior intelligence, but they lived in houses which were old, and too large by modern standards, and gradually their parents moved out to modern houses in more desirable areas. As they moved, their houses became tenements and the character of the school population gradually changed, revealing an increasingly lower level of intelligence. This eventually left a bewildered staff of teachers, uncertain whether to blame themselves or the children for an incompleted syllabus and consistently poor work.

Temperamental differences also will often trick teachers into fallacious judgments about children. For example, two young children, A and B, were examined in a school, and A was shown to be appreciably brighter than B. The headmistress queried this result, suggesting that their names might have been confused because in her opinion B was the brighter one. Further discussion showed that B was of a lively disposition, and was always ready to answer any question asked in oral lessons, while A was retiring and did not do himself justice under examination conditions. It was also admitted that B always produced disappointing written work. The subsequent educational history of the two children amply justified the test result.

### **Test Standardization: Illustration from Spelling Test.**

We thus arrive at a position where we see that subjective factors and special environmental conditions must as far as possible be discounted if an adequate yard-stick or measuring unit is to be devised. A spelling test may be used as an illustration of how this is done, as in this case we are not so readily involved in disagreement about what exactly is being measured as we might be in the case of an intelligence test. A child spells a word correctly or incorrectly and the marking is objective; he is either right or wrong. A list of words can therefore be compiled and given to children to spell, and we might say a child is a good speller or a bad speller according to



what he scores on this test. But suppose different people compile lists. One teacher may have a predilection for words like "catarrh" and "mausoleum", which the children must know before he will consider them good spellers. In the bad old days of "payment by results" in our county schools the Board of Education used to insist that the children should be able to spell all the words they could read, and there are still teachers who, being themselves excellent visualizers, believe that children should be able to spell what they read, or, more modestly, believe that reading improves children's spelling. This, however, except for certain psychological types, is not true. If we elect to take a list of words from a book for the children to spell, we are faced with the questions "Which words?" "Which book?" and we still have not dispensed with the subjective element. The difficulty can be overcome by referring the question to the children themselves. Irrespective of what children "ought" to be able to spell we must ask ourselves what they in fact *can* spell. Lists of words of *apparently* increasing order of difficulty can be given to children, and from the results they achieve, the test deviser's errors of judgment can be detected and eliminated, and a list produced which *children themselves* find to be in increasing order of difficulty. Since the intelligence of the children and the amount of time given to spelling vary from one school to another, these divergencies can be overcome by taking the average of all schools. The labour of such an undertaking would be tremendous, and in reality not worth the effort, since a very close approximation can be obtained by taking a fair number of representative schools and finding the average scores from them.

### Age Factor.

So far, this discussion has left out of account the fact that, within certain limits, intelligence and skill develop with age. Reverting to the idea of compiling a list of words which *children themselves* find increasingly difficult to spell, we see that, provided we still have in mind the "average" child, we



must allow for the fact that young children can spell fewer words than older ones. This "average" child is a representative pupil, and is found (to a near approximation) by taking the average child from a number of schools selected to be representative of schools in general.

We could thus obtain spelling tests for infant, junior, and secondary school children, but in point of fact it is more convenient to have one test, especially in spelling, for all children, divided into age-groups. Thus there are words suitable for 5-year-olds, 6-year-olds, and so on, up to the age of 13 or 14 years.

### Brevity of a Test.

A teacher seeing a standardized spelling test for the first time may be surprised to find that, as in Burt's *Graded Vocabulary Test*,<sup>1</sup> there are only ten words for each year, and he may feel that so few words will be inadequate to test a child's spelling ability. Two comments can be made on this. In the first place a deviser of a test has to consider the amount of time that can justifiably be spent on testing. Indeed some teachers reject psychological tests, solely on the ground that they take up too much time. It would be wrong, of course, to make this question of the economy of time the sole criterion in constructing a test, for it could, presumably, be so brief as to be worthless.

### Selection and Placing of Words on the Scale.

It must be borne in mind, however, that Burt's ten words for each age-group are representative words. That is to say they are finally chosen, after many words have been tried out on the representative children, so that they are as good as, and, for certain reasons, probably better than, other words appropriate for children of the particular age under consideration. A warning must here be added that the words in a test have no greater *intrinsic* value than other words. I discovered that,

<sup>1</sup> *Mental and Scholastic Tests*, second edition, page 382.

after hearing me lecture, a teacher was advising a colleague to teach the words in Burt's Spelling Test. She had apparently decided that these words were most important in themselves. It should be obvious that, if the words in a test are specially taught to a class, the test is of no further use for those particular children.

In deciding which words are to be considered representative for an age-group, the most important point to remember is that half the children of the age-group must be able to spell them correctly. At first sight this might seem a rather small proportion. In the early days of standardizing tests, various standards were adopted for assigning an item to a particular year. Readers wishing a clear exposition of the need to make this figure 50% rather than 75 or 80% for the proportion of children correct, are referred to Burt's *Mental and Scholastic Tests*.<sup>1</sup> Briefly, we may say there are good spellers and bad spellers of varying degrees, while the majority have average ability in spelling. If we choose words that almost all the children of the age we are considering can spell, they will be too easy for the best spellers. If we choose words that very few can spell, they will be much too difficult for the worst. If we choose words which half the children can spell, we have then *representative* words for the age-group.

### Use of the Test.

Finally, in the actual use of the test, those children who are not able to spell the words assigned to a given year are given words chosen for younger children, while those who find them too easy are given words suitable for older children. In practice it is not found that children have a high proportion of words right up to the list for a certain year, and then have none right. On the contrary, their failures are spread over one or more years. They are given credit for *all* the words on the list which they can spell. In Burt's test, since there are ten

<sup>1</sup> Page 150.



words per year, each word is worth approximately one-tenth of a year of spelling age, so that the number of words correctly spelt, divided by ten, gives the spelling age. Since, however, children do not spell at all under the age of 5 years, this time is added on to the score, so that "spelling age" is comparable with chronological age.

It should now be obvious why such a test is called a "standardized test". It is constructed to fit certain standard conditions, and if correctly administered will produce a satisfactory result; it makes a reasonably valid comparison possible between one child and the hypothetical "average" child, between one child and other children, and between different groups of children.

### Standardized Reading Test. Fulfilling the Conditions of Standardization.

Burt has also a reading (accuracy) test<sup>1</sup> based on the same principles, which is very useful for younger children and backward older ones. It has been rearranged by P. E. Vernon, and published with an equivalent or parallel test in one of the *Publications of the Scottish Council for Research in Education*.<sup>2</sup> This exceedingly cheap publication not only provides two excellent word recognition tests, but also gives an interesting account of the process of standardization, and contains information which should convince the open-minded as to the high validity and reliability of the test.

There is one fundamental consideration which must always be borne in mind when using a standardized test, namely that as much care must be exercised in giving the test as was given by the author of the test to its standardization. For example, Burt's Spelling Test was standardized by dictating the list of words. Each word was pronounced separately, and repeated if necessary, as much time was allowed as the children needed, and the children wrote their spellings. In giving the

<sup>1</sup> *Mental and Scholastic Tests*, page 367.

<sup>2</sup> XII. *The Standardization of a Graded Word Reading Test*.



test these conditions must also be repeated. The test would involve a different order of difficulty if, for instance, the child were asked to spell the words orally, or if he were hurried through it, or denied the repetition of the word which he had not grasped at a first hearing.

### Timed Tests.

Some tests are timed, and these timed tests frequently cause a good deal of anxiety among teachers who have never had an opportunity of understanding completely standardization in mental tests. The standard conditions are also most frequently violated in the case of timed tests, especially where the test contains sub-tests with brief timings. Teachers will often say that the time allowed is ridiculously short, and that hardly any of the children have time to finish. The fact is, however, that on timed tests the majority of the children should not be able to finish. The time limit is set to increase the difficulty, the assumption being that the "better" child will achieve more in the set time than the "poor" child.

### "Headroom."

This point also introduces the question of the "headroom" needed in test construction. The purpose of testing is to compare children. If they all finish a test successfully they all have the same scores, and therefore cannot be divided into better and worse. Of course, if a test is set on a definite and limited piece of work which has had to be learned and the task has been of only moderate difficulty, many children will have full marks. Such an *ad hoc* test is useful, but is of no administrative value, and is of no use in assessing the broad progress being made by a child or group of children. Further, such a test is completely out of place when innate intelligence is being measured, for in that case the "learned" response is avoided so far as it is in the power of the test constructor to do so. Therefore all good tests, if they are applied within the conditions laid down in the standardization, should have

"headroom", that is, the child should always be able to do a little more at the test if he had a little more time, or a little more ability.

Thus when the teacher feels the test does not allow enough time he must remember that the original "standard population" of children had a similar amount of time, and to allow more will give his children higher scores, and therefore make them appear better than they are. To be soft-hearted and allow an extra half-minute on a test of three or four minutes' standard duration is really a sort of cheating—self-deception if the test is an internal affair and nothing beyond the teacher's interest hangs on it, dishonesty if the test is administered for an outside authority.

### **Simplification of Measurement.**

Another principle in measurement that is well worth understanding is that tests, as far as possible, aim at measuring one thing at a time. Thus in arithmetic we have tests of mechanical skill and skill in problems. In reading there are tests of word recognition, speed of reading, and comprehension. It is a very common practice to test the reading of primary school children by hearing them read, and awarding a mark subjectively, that is to say, on the teacher's opinion of what the child's effort is worth. Successive children may be asked to read successive paragraphs from a book. Here at once the first principle of scientific testing is violated, for the tasks set for different children are inevitably of varying degrees of difficulty, and yet they are presumably being compared and awarded marks on the same scale. Again, what is usually gained by this method is a *general* impression of the child's reading. There may be great gaps in the order of difficulty of the words read, many easy words and one or two relatively difficult ones. One or two of these "hurdles" may upset one child much more than another, and give a false picture of his real reading ability, especially if the teacher sets great store by fluency. Such a method has two further defects. No one



who knows children intimately can avoid what the psychologist has called the "halo effect". If a child is good in other subjects, and has generally a satisfactory record, we unconsciously presume that his performance will be good. Similarly, if he is careless and untidy in his appearance, or with his books, or if he has brothers or sisters who have had a bad educational record, we unconsciously bring to bear comparisons which influence our judgments. Another disadvantage of the subjective method is that it very rarely gives enough "spread" to the scores. If the teacher is marking out of 10 he may find the effective scale is only 5 or 6 marks, for possibly no child will be awarded less than 4, or, if the teacher is severe, no one may reach above 8 or 9. A well-constructed *objective* test, Schonell's "My Dog"<sup>1</sup> test for instance, would avoid these defects, and separately assess speed, accuracy, and comprehension. Moreover, the results would have the not inconsiderable advantage of making the children tested comparable with a "standard" child population.

### Intelligence Tests and Attainment Tests.

Mention has already been made in Chapter I of intelligence tests and attainment tests. In an earlier part of this chapter the process of standardization was described with reference to a spelling test, which was chosen as the simplest illustration of the general principles involved. The argument then proceeded on the basis of standardizing *attainment* tests. Those remarks, however, are applicable to testing in general, and we must now consider *intelligence* tests also. At the risk of being tedious I would again call attention to the difference between intelligence and attainment, for experience has shown that with the uninitiated this differentiation is very necessary. Intelligence tests measure an inborn characteristic, while attainment tests measure acquired achievements. One final remark about standardizing attainment tests should first be made. Most

<sup>1</sup> *Backwardness in the Basic Subjects*, page 522.



attainment tests in this country are based on "age performance", that is to say, the *norms* or attainment ages allocated to the various scores or marks obtained are arrived at from the score obtained by the average (or more accurately the median) child in the age-group. American tests are frequently based on "grades" which are similar to the old "standards" in this country, a term in use when promotion was based on performance rather than age.

Intelligence tests also are most frequently based on age performance, and yield a "mental age". Thus if a child of 10 has as many items right in an intelligence test as the average 10-year-old his mental age is also 10. If he obtains the score of the average 8-year-old his mental age is 8, and he is considered two years retarded mentally. This last statement needs a little amplification to make its meaning quite clear. Although it is the aim of this book to present the subject of mental testing in as simple a way as possible, and not to enter into theoretical discussion, some reference must be made to the nature of intelligence.

### Narrow Scope of Intelligence Tests.

So far as the ordinary standardized intelligence test goes, especially the verbal one, what is measured is largely that which enables a child to succeed in ordinary school work. The amount of agreement or the correlation<sup>1</sup> between intelligence test results and teachers' estimates of children's intelligence is quite considerable, as is also the correlation between intelligence tests and tests in the 3 R's. In *Mental and Scholastic Tests*<sup>2</sup> Burt describes intelligence as "inborn all-round mental ability" and in *The Backward Child*<sup>3</sup> as "innate, all-round, intellectual efficiency". I personally prefer the second definition, for intelligence tests give little scope for the exercise of intuition and the expression of feeling so often needed in the

<sup>1</sup> The amount of agreement between variables is usually expressed as a correlation coefficient.

<sup>2</sup> Page 211 (second edition).

<sup>3</sup> Page 11.

exercise of "common sense". Both intuition and feeling<sup>1</sup> are "mental" abilities of great value to the total personality in the art of living, but intelligence tests deal only with the thinking side of life, and in the main this thinking takes the form of words. It is because they have failed to realize the relatively narrow scope of intelligence tests that many teachers (and others) have given them a good deal of inappropriate criticism.

There is the further question of the manifestation or use of intelligence in specific ways, as, for instance, in a mathematical or in a mechanical way. These "specific abilities" are known to exist, indeed are easily recognized as being possessed by certain gifted people, but in actual fact they are rarely in evidence in young children. A growing body of evidence is being assembled which points to the fact that these "specific abilities" tend to begin to mature in adolescence. The possession of such an ability may be due partly to the canalizing of intellectual energy in a certain direction through interest and opportunity. From the point of view of the teacher, however, especially the primary school teacher, the most important thing is to be able to measure the "innate all-round intellectual efficiency" and this is done by means of an intelligence test. We are thus enabled to estimate what a child may be expected to achieve in several directions, especially in the 3 R's. Intelligence tests will also prove very efficient in selecting those children most likely to succeed in a grammar school. (The question of tests which will prognosticate success along certain lines at the secondary stage is dealt with elsewhere.)

### Standardization of Intelligence Tests.

Some comments on the standardization of intelligence tests can now be made. Some of these are standardized very much as a spelling test is standardized, that is, item by item. The Terman Merrill test, an individual test used by clinical psychologists, has been dealt with in this way. For instance,

<sup>1</sup> Feeling is here used in the sense used by Jung. See his *Psychological Types*, page 543.



the ability to reproduce a circle, a square, and a diamond to certain well-defined standards of accuracy, has been assigned to the third, fifth, and seventh years respectively. As with the spelling test the child's scores from all parts of the test are added together to give a total number equal to the mental age, and the scale is a composite one, proceeding from 2 years to the superior adult level. The type of test teachers will usually find convenient to use, however, is a group test. Here the test is standardized for a narrow age-group, say 8 to 11 years or 10 to 12 years, and the mental age is obtained from the total score of, say, 75 or 100 items. These test items are contained in a specially prepared test booklet and often require nothing more than the underlining or the crossing out of a word. This makes the marking of the test objective, which, as has been stressed before, is essential for standardization. Various arrangements of test items are adopted in different tests, each claiming some advantage over the others, such as ease of administration, reduction of fatigue, and so on. In some tests there are sub-tests separately timed; in others different types of test item are judiciously mingled and a total time set. Whatever test is used the standard instructions *must* be meticulously observed if the results are to be of any value.

### Criticisms of Intelligence Tests.

Very few tests are produced which a teacher would be prepared to approve without qualification. Here and there a test item may seem to be inappropriate to the experience of the children in a certain district. A level-headed psychologist, however, would be only too ready to admit that the perfect intelligence test has not yet been devised. Nevertheless, a well-standardized test, in spite of minor imperfections, will do its work quite well. It should be remembered, in the first place, that one unsatisfactory item out of a hundred is not going to make a great difference to the results achieved, and in the second place, that an unsatisfactory item in most cases will be equally



unsatisfactory for the children forming the "standard population" when the standard scores are worked out, and again will not unduly invalidate the test. Critics of tests should also guard against bringing their own intelligence to bear on a test designed for a child, and assuming that it will measure appropriately in both cases.

Two illustrations of different types of misunderstanding might be mentioned. A "non-verbal" test had one item in a "classification" sub-test which involved selecting from a row of geometrical figures those which were of the same type as three given figures. It was first necessary to decide what was similar about the three figures, which consisted of a straight line and arcs, intersecting or touching. An adult with some knowledge of geometry finds this much more difficult than the unsophisticated child, for the lines simply form closed or open figures. The three given figures are closed, and it is necessary to indicate which in the rest of the row are closed. The adult is puzzled and looks for more difficult relationships between the lines, whereas the child sees the point at once. Another illustration of the inappropriateness of an adult judgment on the effectiveness or "validity" of tests for children is provided by the "substitution" test. In one form of this type of test there is a row of pictures—a clock, a hat, a football, etc.—and beneath them are drawn, in the order stated, a dot, a vertical stroke, a ring, etc. The child's task is to note the relationship between the picture and the sign beneath and in a given time to place the appropriate signs beneath a number of pictures arranged at random. In this test for young children the manual effort of marking the signs is slight in relation to the intellectual effort involved in thinking out the correct relationship between sign and picture, while the situation is reversed for a grown-up, for whom the test becomes one of speed rather than of intelligence.

Thus the teacher using a test may be able to criticize the test construction, but on the one hand these criticisms may be of little weight compared with the value of the test, and on the other hand they may actually be ill-founded, owing to the

critic's lack of knowledge concerning the mental process which the test is actually exploring. Most of the tests at present on the market are quite sound. They are both reasonably valid and reliable, valid because they test what they set out to test, and reliable because they can be given by various people to the same child at different times, and yield more or less the same result.

### Validity and Reliability.

The validity and reliability of a test can be measured statistically, and in the case of many tests their coefficients of validity and reliability are shown in the explanatory handbook which accompanies them. These coefficients are expressed as decimal figures: for example, .6 would be an unsatisfactory figure; .9 would be acceptable; 1.0 is the unattainable ideal. It is very desirable that all tests should have their "hall-mark" fixed to them in this way.

A test is *valid* if it measures what it sets out to measure, and it is *reliable* if it can be used on successive occasions, with the same children, to give results which agree with one another. No test can be expected to give exactly the same result on two successive occasions, as the response of children will vary somewhat from day to day, but a good test should place the children approximately in the same order when it is readministered. If the test gave exactly the same order, the reliability coefficient would be 1.0, but this degree of reliability is never reached. Complete absence of agreement in the two results would give a reliability coefficient of 0. A coefficient of .95 would be very satisfactory. As regards *validity*, two tests claimed to measure general intelligence should give very similar results if administered to the same group of children. If they give widely dissimilar results, one (or both) must clearly be failing to measure general intelligence, and will thus have poor validity. If, however, two different tests give almost identical results, they must obviously be measuring the same thing, and they will have high coefficients of validity.



The actual calculation of validity or reliability coefficients is a matter of mathematics and does not concern us here.

### Coaching.

It goes without saying that published tests which are not standardized are useless as measures of intelligence. They are used sometimes for coaching children, a practice which cannot be too strongly condemned. In the old days of competition for a very few grammar school places the Scholarship Examination was a hurdle surmountable only by means of extensive "cramming". The system was bad in many ways, but at least when children were "crammed" in a certain subject they did as a rule improve in that subject. "Cramming", however, can do nothing to make a child more intelligent, though it may make him appear more intelligent on the examination day. Teachers have a great responsibility in this matter, for if they coach the child so that he gains a higher mark in an intelligence test than he deserves, they may be responsible for subjecting him to great strain or even failure, if he proceeds to a higher education from which he is unable to profit. They may also effectively keep another much abler child out of a grammar school.

A good deal of research has been carried out on the effects of coaching. Its effects in individual cases are very varied, but it would appear that, on the whole, particularly where there is practice with tests of the same type, the intelligence quotient can be raised. There seems, however, to be a limit to how much can be done in this direction, and the solution seems to lie partly with the teachers and partly with the Local Education Authorities who arrange Common Entrance or Transfer Examinations. The teachers should add to their rules of professional conduct some provision to cover the misuse of intelligence tests; the Local Education Authorities, on the other hand, should arrange that all children about to be tested by means of a certain type of test should be familiarized with that particular type to an equal degree.



### Intelligence Quotient.

Mention must now be made of the concept "I.Q.". The "intelligence quotient" is a technical term which has come into common use, and which it may seem unnecessary to explain. I have encountered so many misconceptions about intelligence quotients, however, some of an obviously crude kind, and others due to lack of technical knowledge about tests, which the average teacher could hardly be expected to possess, that some further explanation may not be inappropriate.

In the first place it should be remembered that the "Q" in "I.Q." stands for quotient and not for quota. It is true that a person's intelligence quotient remains more or less constant throughout his life, which fact may have given rise to the belief that each person has a fixed "quota" of intelligence and that "I.Q." stands for "intelligence quota". But in the conception of I.Q. the idea of quota finds no place; the essential idea is that of ratio—indeed, Burt has used the term "mental ratio" (M.R.) in place of "intelligence quotient". This mental ratio is the ratio between mental age (M.A.) and chronological age (C.A.) and is expressed in the form:

$$\text{intelligence quotient} = \frac{\text{mental age}}{\text{chronological age}} \times 100.$$

The quotient is multiplied by 100 to bring it to a whole number. Thus a child with a chronological age of 8 years 3 months, and a mental age of 6 years 2 months (a very dull child, incidentally) would have an intelligence quotient of 75. The quotient is in fact 74.74 but is given to the nearest whole number. This is only a matter of common sense since, if the intelligence quotient is reckoned on different occasions, it may vary several units up or down, so that decimals are an entirely unjustifiable refinement.

### Constancy of Intelligence Quotient.

Mention has already been made of the constancy of the intelligence quotient. This must now be related with the statement

that it is not uncommon to find that the intelligence quotient does vary. This is another of those points which requires elaborating, for a number of qualifications have to be made to the statement that "the intelligence quotient is constant", if the essential truth it contains is not to be misleading.

In the early days of mental testing it was noticed, in connection with the examination of children for Special Schools for the mentally defective (or educationally subnormal, as we more properly say nowadays), that when a child was re-examined after a period of time, the amount of mental retardation was frequently found to be increased. This was a matter of great interest, for it was thought, and is still thought by the uninformed, that appropriate education must make a child more intelligent, in spite of the evidence of tests showing that children were lagging more and more behind with each year. Thus a child of 8 with a mental age of 6 years, that is to say two years retarded, would probably be found to be two and a half years retarded by the time he was 10 chronologically. This was found to be more or less true, not only for individual children, but for the average of classes of successively older children, so that in a Special School the 7-year-olds might on the average be two years behind in mental age, but the 10-year-olds might be three years or more retarded.

It was also noticed that the brighter the child the less was this falling behind, and that in the case of children above the average in brightness the amount of relative advance actually increased with the years, so that the 8-year-old with a mental age of 10, that is to say, two years advanced, would probably be two and a half years advanced by the time he reached 10 chronologically. The next step, of course, was to appreciate that there is a constant relationship between a child's mental age and his chronological age expressed not by the *differences* between them—for this changes year by year for all but the average child—but by the *ratio* between them, or the *quotient* when one is divided by the other, and, as has already been explained, the formula adopted is to divide mental age by chronological



age and multiply by a hundred to retain the integral form. Thus, selecting a dull, an average, and a bright child, we may construct a *theoretical* table showing the mental ages for each child at 6, 8, and 10 years respectively. By finding the ratio between mental and chronological age we see that the intelligence quotient is the same at each age.

	Chron. age	Mental age	Intell. quotient	Difference between chron. age and mental age in yrs.	
DULL CHILD	6	$4\frac{1}{2}$	75	$-1\frac{1}{2}$	(Mental retardation increases with chron. age.)
	8	6	75	-2	
	10	$7\frac{1}{2}$	75	$-2\frac{1}{2}$	
AVERAGE CHILD	6	6	100	0	(Mental and chron. age remain equal throughout.)
	8	8	100	0	
	10	10	100	0	
BRIGHT CHILD	6	$7\frac{1}{2}$	125	$1\frac{1}{2}$	(Mental advancement increases with chron. age.)
	8	10	125	2	
	10	$12\frac{1}{2}$	125	$2\frac{1}{2}$	

In actual fact these figures are found to be roughly true in the majority of cases, though, as has already been stated, slight variations are common. This is not surprising, since in the giving of tests so many adventitious factors are involved. A child suffering from toothache or a headache on the day of the test is quite likely to be unable to give the complete co-operation which the test presupposes. The advent of a new baby in the child's home, or the sudden removal of a parent to hospital, might easily provide just sufficient distraction to rob the child temporarily of the necessary powers of concentration, and so deprive him of some of his normal output of mental



energy. Nevertheless, tests are so constructed as to be generally interesting and, certainly in the case of individual tests, the examiner is anxious to set the child at ease, and establish in him a sympathetic attitude to the test, so that such distractions are rarely found to cause more than a few points of difference on successive testing. Greater variations are, however, found here and there in the case of extremely unstable children, pre-psychotic children, epileptics, or children suffering from some disease which attacks the central nervous system. A variation of 5 points of intelligence quotient either up or down, however, should not usually excite comment among psychologists who are accustomed to testing children.

The question of the limit of intellectual growth might profitably be considered at this stage. Just as children during adolescence reach a maximum in growth, and attain a height which is never surpassed in later life, so in intelligence they reach a maximum. This does not mean that one cannot become wiser or more experienced as one grows older, but it does mean that one's capacity to deal with a completely new situation involving one's intelligence or "mental efficiency" reaches a maximum during adolescence. The age at which a child reaches this intellectual maximum is usually taken to be 15 years, though it is known that a bright child's intelligence continues to grow beyond this age, and that a dull child's intelligence growth is exhausted before it. There is also evidence that for some little time before this point is reached there is a slowing down. This is not surprising, for Nature is rarely sudden in her decisions.<sup>1</sup>

<sup>1</sup> I have been deliberately conservative here. Fleming in *Adolescence* quotes research showing that growth may continue to 20 or 30 years. On the other hand, Raven, in the *British Journal of Psychology*—General Section, Vol. XXXIX, Part 1, has carried out a research showing that intelligence, which he identifies with capacity to form comparisons and reason by analogy, reaches its maximum somewhere about 14 years, stays relatively constant for about 10 years, and then begins to decline slowly. It would appear that these findings depend to some extent on the test used, which brings us back to theoretical speculations about the nature of intelligence, a topic beyond the scope of this book. The increased length of the average child's school life, however, now makes the solving of this problem of considerable importance to teachers of adolescents.

We are now confronted with a problem regarding the definition of intelligence quotient, for if mental age reaches a maximum in adolescence, and we continue to divide this by the chronological age, the intelligence quotient will fall. Thus a person who was a child of average intelligence on reaching late maturity would have an intelligence quotient equivalent to that of an imbecile. It is therefore necessary, in order to find the intelligence quotient, to take 15 as the divisor whatever chronological age is reached beyond this. In point of fact, a refined individual test like the Terman-Merrill has a scaled system, the divisor becoming less than the chronological age from 13 years onwards, and decreasing more and more until 16 years, after which it remains constant. In this way the gradual tailing off of the growth of intelligence is taken into account.

A little reflection will now show that the devising of an intelligence test for adults (or children in late adolescence) which would measure in terms of intelligence quotient presents something of a problem. So long as intelligence continues to grow no difficulty arises; a child of 8 years may have a mental age of 10 years and from that information we can calculate the intelligence quotient. Similarly, we can easily find the intelligence quotient of a 12-year-old who has a mental age of 15, but when we reach the limit of growth we find that the average of the 17-year-olds, or the 18-year-olds, is always equal to the average of the 15-year-olds. Thus the calculation of an intelligence quotient, at least for those of above average intelligence, seems impossible. There is, however, a solution of this difficulty. Just as the observation of the uniformity of growth of intelligence led to the conception of the idea of "intelligence quotient" or "mental ratio", so another observation of uniformity, this time the uniformity of the *distribution* of intelligence quotients, led to the possibility of awarding to adults intelligence quotients which would be a fair assessment of their level of intelligence.

For instance, it was noticed that there were as many high intelligence quotients as low ones. If, for the sake of argument,



there were 1% of intelligence quotients over 135, there would be found to be 1% below 65; if 25% of intelligence quotients were found to be between 100 and 110, then there would be found also 25% between 100 and 90, and so on. This form of distribution follows a well-known mathematical curve known as the normal curve of distribution, and it is possible, by carefully selecting test items, to make the results of a test given to a large representative group of adults follow this normal curve. It is this procedure also which is usually adopted when a group test is devised for younger children.

When this normal distribution was first empirically determined, it was found that there were many average children and few extremely dull or extremely bright. There is a statistical device, known as the *standard deviation*,<sup>1</sup> for calculating the manner in which scores spread out above or below the central intelligence quotient of 100. Thus if the standard deviation were 10 (points of I.Q.) the intelligence quotients of roughly two-thirds of the group of children would lie between intelligence quotients of 90 and 110 (i.e. 10 points below 100 and 10 points above 100); and if the standard deviation were 15 (points of I.Q.) the same two-thirds of the group of children would have intelligence quotients which would lie between 85 and 115 (i.e. 15 points below 100 and 15 points above 100). The latter obviously shows a wider range. It follows that if the standard deviation of a test is high there will be more high and more low intelligence quotients and fewer average ones, and if the standard deviation is low there will be fewer high and fewer low intelligence quotients, and correspondingly more average ones. When the earlier versions of the Binet type tests were being standardized the standard deviation was found to be 15 points of intelligence quotient, and some tests, for instance the well-known Moray House tests, aim at reproducing this.

<sup>1</sup> It is outside the scope of this book to elaborate this concept, but it is dealt with in Vernon's *Measurement of Abilities* and books on statistical method.



In accordance with the general policy of this book, which is to avoid technicalities as much as possible, I should perhaps have omitted the question of standard deviation, but as it has some bearing on the problem of the constancy of intelligence quotient to which reference has already been made, it seems necessary to give it some consideration at this point. As we have seen, certain adventitious factors may cause greater or smaller (usually smaller) variations in intelligence quotient, but the size of the standard deviation also influences the intelligence quotient. Thus we find that the Terman-Merrill test has a standard deviation of 16.5 points of intelligence quotient, which means that an intelligence quotient of 73 on this test would be equivalent to 75 on the Moray House test. Again, the test with the bigger standard deviation gives the lower intelligence quotient as it also gives the higher intelligence quotient at the other end of the scale: thus a Moray House intelligence quotient of 131 would be equivalent to 134 on the Terman-Merrill. These are only slight differences which become still less as the intelligence quotients approach 100, but with a large standard deviation such as is claimed for certain Cattell tests (27 points of intelligence quotient) the usual intelligence quotient values are completely upset. Moray House intelligence quotients of 75, 90, 110, and 125 become on these tests respectively 56, 82, 118, and 144.

These differences are not trifles. Whether the reader has been able to understand these facts or not, it must be emphasized that the intelligence quotient given by one test is not necessarily the same as the intelligence quotient given by another. It is always necessary to know from what test an intelligence quotient is obtained before we can understand its significance. Because of the variability of intelligence quotient owing to different standard deviations, it is always best to compare children, as far as possible, by means of the same test. At the same time, intelligence quotients from two or more tests cannot be averaged for greater reliability unless the tests have the same standard deviations.

It is in dealing with such technical points as these that the employment by the Local Authority of an educational psychologist is most desirable. In fact, where authorities undertake testing programmes in conjunction with Transfer Examinations and record cards such an appointment is essential. Where these matters are left to untrained administrative staff grave injustices may be done. Authorities have been known to conduct examinations where "raw" attainment scores (i.e. English and arithmetic marks) have been added to intelligence quotients to obtain a final score, while intelligence tests have been sent out for administration without adequate instructions.

### **Percentile Ranks.**

Many psychologists deplore the continued use of the term "intelligence quotient", and would prefer to use the concept of "percentile ranks". This would certainly have many advantages, and as it is a matter of some importance I shall deal with it in Chapter VI, but until psychologists themselves prepare tests with the scores appropriately expressed in percentiles, we cannot dispense with the system which I have outlined in this chapter.



## CHAPTER III

### TYPES OF TEST

Before I begin to describe the selection of tests for specific purposes, it will be advantageous to consider the whole range of tests likely to be of use to teachers. Reference has already been made to the difference between *intelligence* and *attainment* tests. In addition we have to consider diagnostic tests, tests of special aptitudes, and other psychological tests not so readily put into any category.

#### Intelligence Tests.

Burt's definition of intelligence as "innate all round intellectual efficiency" has already been accepted as adequate. Intelligence tests set out to measure this inborn quality, and in their operation obviously must set the mind to work. The quality of what is achieved is always the prime consideration, and in some cases both quality and quantity of mental "work" may be assessed.

#### Individual Tests.

All early tests of intelligence were individual ones; that is to say, one child was examined at a time with a standardized series of questions. Such tests aimed at discovering levels of development reached in matters of practical everyday life. For example, the ability to draw a square and a diamond, to give change, to define words, to reproduce designs from memory, and to detect absurdities in certain statements was tested. Many of the early tests have been discarded in course of time, because, for one reason or another, they have been found un-

suitable. For instance, children in the fifth year were expected to name correctly four primary colours, in the tenth year they were expected to arrange five weights in order, and in the thirteenth year, after being shown three unequal lines, they were asked which of two equal lines was the longer; this was a test of the child's power to resist suggestion. These particular tests have all been given up now, and only those outstandingly successful for the testing of general intelligence have been retained. The test most widely used in Child Guidance Centres at the present time is the Terman-Merrill test. In order to ensure reliable results this test should be administered and interpreted only by an examiner who has had special training in this field as well as a wide experience with all kinds of children. It is therefore not recommended for use by the average teacher.

Professor Valentine, however, has prepared a series of individual tests<sup>1</sup> requiring a minimum of apparatus, which many teachers would find very useful. The examiner must make a thorough study of this test in order to understand its method of calculating the mental age, and it must be exercised on a fairly large number of children before it will produce results that are completely reliable. It does contain some excellent detailed instructions and advice, but it would be advantageous if the teacher could have a short period of training in its application from a psychologist or some other person competent in the administration of individual tests. It has been proved that this test gives a very good prognosis of the educational progress of junior school children,<sup>2</sup> and it should be equally useful in Infant Schools. Other individual tests have come to be called *performance tests*, but these are not tests of general intelligence. They will be dealt with later in this chapter. One great difficulty about individual tests is the time taken in their administration, but for young children, and especially backward children, there is no doubt that they

<sup>1</sup> *Intelligence Tests for Children* (Methuen).

<sup>2</sup> *British Journal of Educational Psychology*, Vol. XIV, Part III.

are more reliable and more valid than other tests, and almost certainly justify the time spent on them. In the case of the Valentine test it is claimed that the average time for testing a child in the middle of the Junior School is about 20 minutes. A longer period may be necessary for a younger child, especially where time may be required to establish a satisfactory test-relation with him before the actual test is begun.

### Pencil and Paper Tests.

In view of the amount of time used in the individual test a method of examining a number of children at the same time has obvious advantages. The most convenient method of doing this is to have a booklet of test items for the child, in which he has to respond in some simple way, such as underlining a word, inserting a tick or cross, or writing a word or two. Where a test is timed it is obviously unwise to introduce much writing, as the test would then be measuring the speed of writing as well as intelligence, and this would vitiate the result. Moreover, the writing might be illegible or the answers ambiguous, so that the underlining method or some equivalent one is generally adopted.

One of the simplest types of test is a test of similarities and opposites, set out in the form:

fall - drop	...	same - opposite
hot - cold	...	same - opposite
old - young	...	same - opposite

Here the child has to underline the word showing the correct relation between the two given words. It is frequently objected that in such a test a child could gain half marks merely by guessing. The test is often improved by adding a third column "unknown" which makes the choice one of three possibilities, or to discourage or discount guessing one mark may be given for each correct answer, and one taken off for each wrong one. In actual fact, investigation has shown that



there is very little guessing, unless children are instructed to guess if they don't know. The test may be given in the form:

fall means the same as: stumble, hurt, drop, break, cry.

This gives a choice of five answers, only one of which is correct, and it reduces considerably the element of chance. As a matter of interest, a test could be answered by "chance", a dice being used to select which of the five given words should be chosen. When this was done for the Cattell Scale IIA test a score of 29 was obtained, which, if applied to a child in the age-group for which the test is intended—11-15 years—would give an intelligence quotient equivalent to that of a mentally defective. Certain items of a test have been found to show intelligence more satisfactorily than others, and devisers of tests aim at combining test items which will produce a test the solution of which will require a very high level of intelligence. Different devisers, however, have different preferences in this matter. Some arrange a series of sub-tests each self-contained and separately timed, such as tests in similarities, opposites, analogies, classification, sentence completion, number series, substitution, inferences, codes, comprehension of instructions, etc. Other devisers arrange items of all types in one test, with an over-all time limit, but it is usually found that the different types are arranged in cycles. Thus in a test of 100 items 1-5, 26-30, 51-55, and 76-80 might be "classification" items, and "opposites" might come as items 6-10, 31-35, 56-60, 81-85, etc. Whatever the plan of construction, the items are graded in order of difficulty so that the child, having successfully tackled the simple initial questions, proceeds to the rest with a feeling of confidence.

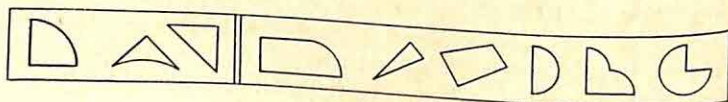
As I do not wish to discuss the nature of intelligence and how it can best be tested, there seems little point in giving examples of all the various types of test items in use. Since the aim of this book is to encourage teachers to use tests and not merely to read about them, mention of one or two tests should suffice to encourage the teacher to seek out tests that will meet

his own particular requirements. As regards the relative merits of the timed sub-test type, and the over-all timed type, advantages are claimed for both, but if they are well standardized and correctly administered, the teacher can use either with fair confidence *provided the test is suitable for the mental ages of the children to be tested.*

### Verbal and Non-Verbal Tests.

Until now it has been implicitly assumed that intelligence tests are verbal tests. Since young children are unable to read, and older backward ones may be poor readers, a "verbal" test (calling for an *understanding* of words) would either fail or be a very imperfect measure for such children. Tests have therefore been devised which dispense with words and use pictures or diagrams instead. For example, one "classification" test shows a row of pictures of toys, into which is inserted the picture of a pair of scissors, and a row of pictures of domestic animals which has among them a picture of wild animals. The one that "does not belong" has to be underlined. It can be seen that whether words or pictures are used makes very little difference to the resulting mental process involved, so long as there is a word corresponding to the picture, or a picture can be drawn to represent the word, and for many tests this is satisfactorily achieved. One criticism which may be made of some test items is that the pictures are not always very good, but uncertainty about one or two items will not seriously upset the test result.

Some items in non-verbal tests are more abstract in character, as, for instance, in a test where the child is required to find what is common to three given geometrical shapes, and thence to pick out two further shapes having the same quality in common from a row of similar shapes, thus:





While such a test is called a non-verbal test it is true that some children talk to themselves while solving such a problem, and so verbalize it. Moreover, the test has a standardized set of instructions which the examiner reads out, and which the child has to understand before he can proceed to do what is expected of him. Truly non-verbal tests have to be self-explanatory or have to be given by dumb-show, as is possible with certain performance tests.<sup>1</sup>

The great value of non-verbal tests is that they dispense with reading. I have come across several boys at different times who have reached secondary school age without being able to read, who have passed through school, being considered virtually uneducable, and who have subsequently been found, by means of non-verbal tests of intelligence, to have intelligence quotients well above average.

Moreover, it should be borne in mind that a child who *can* read may not have *enough* reading ability to do himself justice on the ordinary verbal test. It has been shown <sup>2</sup> that the Moray House test sometimes gives inaccurate intelligence quotients because of lack of reading ability in younger children, so that the test does not function satisfactorily until the child reaches the age of  $9\frac{1}{2}$ . The same research also tentatively suggests that verbal tests may give inaccurate results, where the reading age is below  $9\frac{1}{2}$  years, no matter what the chronological age may be.

### Some Practical Points.

Some tests have been constructed with both non-verbal and verbal items. One test, for instance, has verbal items inserted solely for the brighter children, a wise method provided one is satisfied that the children have no specific reading disabilities which would prevent the full exercise of their

<sup>1</sup> See page 54.

<sup>2</sup> Margaret A. Mellone, *British Journal of Educational Psychology*, Vol. XII, Part 2.



intelligence in the verbal part. This raises an important point about the marking of pencil and paper tests. Where such a test is given to an individual child it is easy enough to watch the child's method of work, but where the test is given to a group this is not possible. It therefore becomes necessary to observe how the child has been working. (It should be pointed out that such observation is not encouraged by the ease and objectivity of marking these tests.) In this type of test, for example where a predominantly non-verbal test has some verbal items added to increase the range of the test, one should not be content with any test result which shows a child scoring well in all the non-verbal items, and failing in the verbal ones. It may be that he has a reading disability which prevents his scoring more, and one would then require to use another test, unless an appropriate reading test showed that no such difficulty could exist.

The child's method of working should be observed too, in order to detect any flaws in the function or suitability of the test. For example, he may not have the necessary "test sophistication", or he may be emotionally disturbed. This lack of test sophistication is most likely to be present in the case of very young children, for whom pencil and paper tests are least suitable. Where, as is often the case, the test is composed of brief sub-tests, it is essential, if the test is to give an accurate result, that the child should be sufficiently mature and adapted to the process of "work" involved to give his best output in the time available. If in certain sub-tests he has failed to make a score, or has scored one point only, the value of the test should be questioned.

Moreover, it might be added as a general principle, if it is found that a child makes a very low score on a test, that little reliance should be placed on the result even if the published norms indicate that the score can be considered equal to a given mental age. If a child has only two or three items correct in a test his intelligence cannot be fairly measured by that score. It is quite possible that he has been doing nothing during most

of the time allotted for the test and he probably needs a simpler or a different type of test.

The great advantage of the paper and pencil tests is that they can be given to a number of children at the same time. They are in fact most usually called "group" tests, though they can, of course, be used equally well for an individual child. It should be pointed out, for the sake of anyone proposing to use group tests for young children, that the size of the group must be severely limited, if the test is to be used efficiently. It is usual for this type of test to contain practice items to familiarize the child with the type of task to be undertaken, these being, of course, included in the test booklet. Most tests for young children are of the timed sub-test type, and there are several practice items to be carried out at the beginning of each sub-test. It is absolutely essential that the child should understand, through these items, what the nature of the test proper is to be. The examiner must make sure that the children are co-operating, and he must therefore be able to see what each child is doing. It is unwise to have more than a dozen young children in the group at once. I have found that the most satisfactory method is to have one child in each desk on either side of a single gangway which can be surveyed by taking no more than a few strides, although even this involves quite a task, when one has to attend to the standardized instructions, time the tests, see that the children do not turn over two pages at once, thus using the wrong test, and at the same time keep them interested, so that they feel the whole thing is more like a game than a frightening exercise. This arrangement is also needed where dull and backward children are tested with non-verbal tests, and of course, with any test and any children, sufficient space must be allowed to avoid copying. A stop-watch becomes almost a necessity for the satisfactory administration of many tests. It should be checked for accuracy from time to time.



### Other Pencil and Paper Tests.

Mention has already been made of Earl's *Tests of Ability for Secondary School Courses* and his "Duplex" ability tests. These provide for the exercise of intelligence in a canalized fashion and test specialized abilities as well as general ability. Other tests are devised to measure the practical or spatial judgment factor, and mechanical aptitude. These are no doubt very important to technical school teachers and administrators, but since there is still controversy<sup>1</sup> about the age of maturing of these factors, and the age at which they can be satisfactorily measured, nothing will be said here about the relevant tests. Some suitable test material is, however, included in Chapter VIII, and more will be found in the symposium mentioned in footnote 1. Moreover, the courses offered by technical schools vary so widely and involve such a variety of skills and abilities that the topic is beyond the scope of this book. The matter is important, however, and interested teachers might profitably pursue the matter further.

The child's curiosity about technical subjects will probably have to be taken into consideration in any ideal scheme of technical tests. There is already some evidence to show that vocational guidance, for instance, when given on the basis of "interest" tests, has been at least as successful as that based on the results of tests of ability. An attempt to deal with this matter in Holland can be studied in the *Mathematical and Technical Tests* of Dr. J. Luming Prak. The test material is available in English (see Chapter VIII).

### Performance Tests.

All tests are, of course, tests of performance, but certain ones, invariably involving the manipulation of material, have come to be called *Performance Tests*. For instance, the *Seguin Form-*

<sup>1</sup> See *Symposium on the Selection of Pupils for Different Types of Secondary Schools*—British Journal of Psychology, Vol. XVII, Parts II and III; Vol. XVIII, Parts I, II, and III; Vol. XIX, Part I.

*board*<sup>1</sup> is a board with ten shapes cut out of it, such as a circle, a star, a cross, a square, and so on. Pieces of wood of the identical shape have to be inserted into holes in the board. The test is standardized, and a measure of intelligence is obtained from the time taken to insert the pieces correctly. This test is quite a useful brief measure of intelligence for nursery and infant school children. For older children it becomes mainly a test of speed and manual dexterity. Though such a test is too brief to give an estimation of intelligence which is anything more than a pointer in itself—it takes at the outside four minutes for three trials—it is useful as a check on other tests, or as a supplement to them. It has a high interest value, is usually regarded as a game, and is useful with relatively uncooperative children. Further, since it is entirely non-verbal it can be used with young children whose speech is unintelligible, or with deaf children. Performance tests are, of course, indispensable for testing deaf children, unless they have acquired enough reading ability to attempt satisfactorily the ordinary verbal group test. Other easily administered performance tests for primary school children are the Porteus "Maze" Tests,<sup>2</sup> Burt's "Drawing" Test,<sup>2</sup> and Goodenough's "Drawing" Test.<sup>3</sup> All three are quickly administered, and no apparatus is required beyond copies of the mazes through which the child has to trace his way. Both Burt and Valentine have ranked this test high as a single performance test, where time is limited. The drawing tests require the child to draw a man, and the result is compared with median samples to arrive at the drawing age. It is most useful as a measure of intellectual development for young children and very dull children. The Goodenough test is more frankly aimed at estimating intelligence and ignores artistic ability. These performance tests are all useful as *supplementary* tests of intelligence.

The *Cube Construction Test* is another performance test.

<sup>1</sup> Described in Cattell's *Guide to Mental Testing* and Gaw's *Industrial Fatigue Research Board Report*, No. 31.

<sup>2</sup> Given in Burt's *Mental and Scholastic Tests*.

<sup>3</sup> *Measurement of Intelligence by Drawing* (F. L. Goodenough).



Wooden cubes, painted only on certain sides, have to be built up into given shapes so that only painted faces show in the completed model. There are three such tests, each being scored by time taken and number of errors made. This test is included, though with different methods of scoring, in both the *Drever-Collins Performance Scale* and the *Alexander Performance Battery*. It is claimed that the latter test measures the special type of ability needed for success at a technical school, and that this ability is measurable as early as 11+.

The following comments can be made about performance tests. Though they measure intelligence and correlate, to some extent, with other intelligence tests, they measure a factor specific to themselves, or common to the battery of tests of which they are a part. They will often pick out a child with a good deal of ability of a special character. Alexander<sup>1</sup> identifies the special ability measured by his own test battery with a "practical factor", but this must not be taken to mean that these tests will necessarily reveal the practical or "handy" person. It seems that such tests measure the ability to make spatial judgments, such as would be needed in technical drawing, design, and certain technical processes, but it does not follow that those well endowed with this ability will also be endowed with the high degree of neuro-muscular co-ordination so often necessary in skilled craftsmanship. It appears necessary to emphasize this point, because many people have used performance tests and, as with other tests, have expected them to yield more information than they are capable of giving.

Performance tests also appear to be influenced by temperamental and emotional factors. Nervous children, and children who expect too much of themselves, show their disability clearly under performance test conditions. Such revelations are of course very useful to those of us who are interested in children first, and test results second, or in tests only in so far

<sup>1</sup> See his *Performance Battery* (Nelson). His point of view with regard to types of ability and "factors" is stated non-technically in *Educational Needs of Democracy*.

as they help our understanding of children. Observation of the child's method of work in performance tests is often illuminating. One child must handle the test material and proceed by trial and error; another is more reflective and thinks before he acts. This is probably in part a temperamental difference, and it is not necessarily true that the one form of ability is of a higher order than another, though the child whose trial and error attempts are of an inferior order makes his inferiority very obvious.

To obtain the best results from performance tests a good deal of experience in using them and observing children's reactions is necessary, and experience can be gained only through time. In an age when so many people try to pack several life-times into one, and when it often seems more important to count many heads than to examine one, this condition may be considered a drawback. The problem, of much wider significance than its mere application to psychological testing, is one which each one of us has to solve for himself. Nevertheless, if we are to avoid becoming entirely mechanical and closing our eyes to all anomalies, we shall be forced at times to feel dissatisfied with the straightforward pencil and paper test for certain children. Where there are doubts in our minds as to the result of a test it is always wise to pursue the matter further, and for further investigation performance tests are often of great value. If lack of time or apparatus is likely to cause insuperable difficulties, the Local Authority's educational psychologist should be consulted, or the child should be referred to the Child Guidance Centre. It is a common misconception that such centres handle only very serious problems.

### **Musical Ability.**

A recently introduced series of tests of musical ability and appreciation by Dr. H. Wing has been most successful. These tests are dealt with under "Test Material" and "Bibliography" in Chapter VIII, and will no doubt be very useful to teachers of music and others interested in gifted children.



### Attainment Tests.

It has already been explained that attainment tests are tests of acquired skill in school subjects, and here it may once again be emphasized that, although they yield attainment ages derived from children's actual performance in representative schools, the final result is to some extent governed by what teachers expect from children. Vernon<sup>1</sup> discovered that Scottish children of 6 to 8 years are approximately six months in advance of English children in reading, a difference which increases to twelve months by 10 to 12 years, but the lead is lost again at 13 years. He attributes this to the Scottish teachers' drive to put the children through the "Control" Examination.

### *Reading.*

We have already seen that different aspects of reading can be tested. The most useful are tests of word recognition and of comprehension. The former are individual tests while the latter may be given either to groups or individuals. Other tests measure speed and accuracy, or a combination of speed, accuracy, and comprehension. Some tests give a combined measure for these qualities, or are scored in such a way as to give a separate measure for each. There are also tests of vocabulary, and tests which investigate various aspects of language development. Some of the tests developed by A. F. Watts,<sup>2</sup> while providing measures of attainment, also analyse the child's contact with, and knowledge of, the world of people and things. In this they go beyond the ordinary attainment test, but should assist the teacher to gain insight into children's concept formation, and to arrange and grade work to suit their level of development.

### *Spelling.*

The standardization of a spelling test has been described in Chapter II. Such a test, of course, takes no account of the

<sup>1</sup> *Standardization of a Graded Word Reading Test.*

<sup>2</sup> Published in *The Language and Mental Development of Children.*

context of the words, which deficiency is overcome by the standardized dictation test.

There are several forms of spelling test: some give several alternative spellings for each word, and the child has to underline the correct one; others give words in either a correct or incorrect form, and the child is required to mark the correct and incorrect forms with a tick or cross; but the type of test which requires the words to be given orally, and the child to write the words, is undoubtedly the best. Spelling involves aural, visual, and kinæsthetic aspects under normal conditions of writing, and children who are not good visualizers are penalized by the first types here mentioned.

### *Other English Tests.*

Reference has already been made to Watts's language tests under "Reading". Chapter VIII contains references to tests of grammar, "Construction", and composition. Some of these are undoubtedly of value to teachers of English.

### *Arithmetic.*

While standardized "composite" arithmetic tests may be valuable for special purposes, teachers will usually find it better to use separate tests for "mechanical", and what has variously been called "problem", "reasoning", or "knowledge of methods" arithmetic. It is found that some "mechanical" tests deal solely with the "four rules", while others include decimals, fractions, simple interest, mensuration, etc. Both these "mechanical" types have their uses, but no doubt the simpler "four rules" test should have priority, since children will not understand decimals or capacity measures, for example, if they cannot deal with integers. Besides "mechanical" and "problem" tests there are also standardized "mental" arithmetic tests. A "mental" arithmetic test is differentiated from other tests in that the "working" has to be done mentally, but the final answers may be given orally or written down. Questions can be asked orally or in print. As



tests of mental arithmetic are tests of speed and accuracy, one which is given orally to a class by the teacher may prevent a slow child from doing himself justice, because the teacher may be always moving on to the next item before he finishes the one he is doing. For this reason such tests usually sacrifice the element of speed, and allow unlimited time, accuracy only being considered. Even so, in giving such a test to a class, it is difficult to decide at what rate to give the test in order to be fair to all. This difficulty does not, of course, arise when children can read, and the test can be given in printed form, with a time limit. The oral form of the test is useful for very young children, and for some backward children it may be the only suitable standardized arithmetic test which can be used. In such cases it should be given individually.

### *Handwriting.*

Standardized methods of measuring handwriting have been devised both for speed and quality. Speed is measured by the number of letters written in a given time, and quality is measured by comparing a child's reproduction of a set piece with *representative* samples of handwriting obtained from children of each age-group. Thus a "writing age" can be determined for both speed and accuracy.

For script-writing Ballard has abandoned the age performance principle, no doubt because most schools change over to cursive writing at the junior stage, and he has compiled an arbitrary scale of ten samples with marks awarded from 1 to 10.

In addition to these standardized tests there is Burt's schedule for the "Analysis of Quality of Handwriting".<sup>1</sup> With the aid of this schedule teachers may draw up their own scales with more samples than Burt is able to provide. Some of the factors he lists are size, slant, line (i.e. pressure), alignment, and spacing—the elements of good handwriting—and he puts forward some cogent arguments in favour of his schedule.

<sup>1</sup> *Mental and Scholastic Tests.*

*Graphology.*

The findings of graphologists should be taken into account, however, in considering handwriting. Graphology is in fact, as far as adults are concerned, a kind of "test" of character and personality. Children, too, quite early begin to show tendencies in their handwriting which are true expressions of themselves. For instance, the restless and excitable are prone to write unevenly, while individuality quickly begins to show in slightly original deviations from the school model. Graphology is a difficult subject and one may even require some psychological insight to be able to apply it usefully. Nevertheless some acquaintance with the work of such graphologists as Saudek<sup>1</sup> and Jacoby<sup>1</sup> would indicate that there are limits to the process of making children into "good" handwriters. What at first sight might be called "poor" handwriting may be due to the strong impulses which actuate a child, and which, if tolerably well disciplined, may be connected with strong character. On the other hand it may be related to motor incoordinations arising from chorea (St. Vitus's dance) or from emotional instability. Incidentally, the timid, inhibited, and depressed child who produces tiny cramped writing, often sagging under the line, is more likely to improve his handwriting by general encouragement and emotional support, by efforts to make him feel that he "is somebody", than by constant exhortations to make his writing larger or to be more tidy.

*Tests in Other Subjects.*

Reference is made in Chapter VIII to one or two tests in other subjects, geography, history, chemistry, etc. Very little has been done in the way of standardizing such tests in this country, but there are many American tests available in many subjects, ranging from biology and economics to shorthand and typing. Teachers who could make use of such tests might

<sup>1</sup> See Bibliography.



well attempt to devise their own <sup>1</sup> "new-type" tests. A little reflection will show that an almost universal standardization, such as is possible for intelligence tests and attainment tests in the three R's for younger children, is almost impossible for other subjects. The standardization will of necessity have to refer to type of school, level of general intelligence, and years of study. Nevertheless, for the purpose of comparing one form with another, one "year" with previous "years", as well as for testing acquired knowledge and measuring progress, such tests are very useful. For rapid and comprehensive surveys the "new-type" examination, involving as it does only one word answers or selective underlining, whether standardized or not, is a teaching tool of great value, and deserves far more recognition than it seems to receive at the present time.

*Care should be taken not to confuse subject ability tests with attainment tests. Thus Earle's "Test of Ability in Algebra" is a special type of intelligence test.*

### Sensory and Sensory-Motor Tests.

A book of tests, designed to help teachers to a more thorough understanding of children, should not omit to refer to certain useful tests which are available, either for the teacher's own use or on his behalf by psychologist or school medical officer. Some of these, such as tests for high-frequency deafness or handedness and eye dominance, may be vital in the differential diagnosis of certain educational disabilities.

#### *Right- or Left-Handedness: Eye Dominance.*

This is a subject of considerable importance to teachers, and is dealt with at great length by Burt in *The Backward Child*. He there shows that left-handedness is not the definite characteristic it is often supposed to be, and that merely to judge by the hand used for writing may be quite misleading. He gives a comprehensive series of tests, discusses training the left-handed, and makes practical suggestions. He also discusses

<sup>1</sup> See Ballard's *New Examiner* or Vernon's *Measurement of Abilities*.

mirror-writing and the question of handedness and speech disturbances, and shows that the latter are more frequently found in mixed types, i.e. children who are right-handed for some tasks and left-handed for others, and especially where the dominant foot or eye is opposite to the dominant hand. He also gives useful tests for discovering the preferred or dominant eye.

The reader interested in a further discussion of the relation of handedness and eyedness to reading and writing difficulties, especially orientation difficulties (i.e. left to right or right to left perception and execution of letters and words), is referred to Schonell's *Backwardness in the Basic Subjects*.

### *Hearing.*

Defects of hearing are frequently serious impediments to progress in school work. In *The Backward Child* various tests are described which can be administered by the teacher, but what is probably the most useful test of hearing-for-speech-sounds under natural conditions is the "voice test" expounded by Dr. M. D. Sheridan in her book, *The Child's Hearing for Speech*. On the other hand, in the above-mentioned book by Burt there are some practical comments on the use and limitations of the gramophone audiometer, and a discussion of the function of the pure-tone audiometer, with its relation to high-note (or high-frequency) deafness. The implications of high-frequency deafness in the diagnosis of reading and spelling disabilities are discussed in Chapter IV. Being comparable in some ways to colour-blindness, high-frequency deafness is also an important factor in certain speech defects. The pure-tone audiometer is indispensable for the diagnosis of this condition, and should be available in every area. It is an expensive scientific instrument, suitable only for the use of the expert, school medical officer, psychologist, or teacher of the deaf, but teachers should be aware of the value of audiometric examination, and should be entitled to ask for this examination where high-frequency deafness is suspected (see Chapter IV).



*Vision.*

Sight is probably the sense which plays the greatest part in school work, and the possibility of defective vision must always be kept in mind where backwardness exists. Burt found marked errors of refraction in one-fifth of his backward cases, and slight defects in many more. Though accurate sight testing is the specialist's job, and involves dark room examination of the eye, there is no reason why the teacher should not make rough tests where he suspects defective acuity. There may be long gaps between routine medical inspections, but in any case a thorough eye test is not usually carried out unless the parent or teacher has reason to suspect serious defect, so that the teacher who has acquainted himself with types of defect and some simple tests can often take early steps to have the defect attended to by the expert. In *The Backward Child* Burt gives a very adequate description of defects and a discussion of their relationship to classroom conditions (blackboard and books). He also describes such tests as are suitable for the teacher's use. Any teacher who uses these tests should not overlook the fact that every sensory test is also a psychological test—a fact not always taken into account even by medical specialists—and should pay due attention to Burt's comments under this head. These auditory and visual acuity tests are also dealt with by Schonell in *Backwardness in the Basic Subjects*, although he does not mention the pure-tone audiometer.

*Colour Blindness.*

This defect appears to have no connection with retardation in the three R's, but it can be a considerable inconvenience in art and needlework. It is certainly of occupational significance, and since Burt found 3.1% of boys to be colour blind, a test of colour blindness may be very useful in considering any advice on careers where colour vision is important. Teachers in a "railway town" no doubt will find significance in these remarks, as this defect is a bar to engine-driving. Only .2%

of girls of Burt's sample were colour blind, but here again, for the small proportion of girls likely to be involved in an occupation requiring colour judgments, a colour blindness test would be useful.

### Interest Tests.

It is obvious that a child's interests are very important, but they are also in many cases ephemeral, and to some extent depend on environment and opportunity. Schonell (in *Backwardness in the Basic Subjects*) has provided a useful schedule for recording children's interests, and Cattell (in *A Guide to Mental Testing*) has devised an interest test for 14-year-olds. At this age interests are becoming more canalized and specific.

### Character and Personality Tests.

Most psychologists would agree that human nature is so complex and fluid that it will not yield to any simple quantitative expression. Diagnosis of character and personality is as much an art as a science, and tests in this field have much less validity and reliability than tests of cognitive abilities. Nevertheless, a number of tests are included in Chapter VIII with which the teacher may care to experiment. Perhaps the greatest danger in using some of these tests is that the experimenter is apt to project his own psychology into the test, and make it yield more than it is capable of giving. The wise student of human nature realizes how little he knows of his own mind and the minds of others, and is careful not to make sweeping generalizations. Therefore anyone who sets personality tests should avoid prejudice at all costs. If a test yields a result contrary to his expectations, he should not accept it without question; tests are not always right, and even the best of them sometimes miscarry; on the other hand, human judgment is sometimes at fault. This type of test is often of most value, not for the clear composite picture given but for the valuable sidelights revealed, which are often surprising, and which enable the examiner to see the child as an individual.



A test such as Valentine's "Mental Types as Revealed by Associations and Descriptions", simple though it is, can be most illuminating, and though in using it one may too readily place a child in a "class" or "type" on too little evidence, the actual insight gained compensates for any such error. Similarly Cattell's "F" (Fluency of Association) Test aims at obtaining a quantitative measure of temperament, but in many cases the material produced out of the child's mind contributes much more to an understanding of the child than the numerical score.

Word association tests aim at discovering the child's emotional complexes. It is doubtful whether the teacher can make any use of these tests from a diagnostic point of view, but the teacher who is a student of human nature would do well to experiment with them if only in order to become aware of the strength of emotional blockages and resistances. It may be that much of the test material given in Chapter VIII under the heading "Interest and Personality Tests" will be of great value in revealing anxieties, conflicts, emotional isolation, etc. Most teachers do not realize that the "too good" child is often a psychological problem. Child Guidance Centres are apt to find that a considerable proportion of their cases have much "nuisance value" to adults. This is not surprising, as such children demand attention. The inhibited child, the solitary child, and the depressed child, however, often have as much or more need of psychological help, and some of these tests or inventories may be of use in focusing attention on these types.

### Rating Scales.

Burt has produced some evidence to show that "for intellectual abilities, and particularly for special aptitudes, the 'reliability' of teachers' assessments is far lower than that of tests; for educational attainments it is about equal; for moral qualities somewhat better". If this is so it would seem that some mention should be made of rating scales in this chapter,

though they are not "tests". A rating scale is a device for assisting the teacher to make an estimate of the degree to which a child possesses certain character or personality traits. For instance, Schonell's "Rating Scale for Temperament" examines a number of traits, such as self-confidence, persistence, assertiveness, etc. Each trait has five categories. For example, that of "self-confidence" ranges from "extremely confident amounting almost to 'cocksureness'" at one end, to "extreme lack of confidence, over-dependent, declines responsibility" at the other. The teacher must mark where he thinks the child should be placed on the scale, or if he thinks fit he may place the child at an intermediate point between two adjacent categories. Additional information is given with the scale to guide the teacher in its use, and a short questionnaire is added to focus attention on the child at play and at home, so that a one-sided classroom judgment is not made.

Rating scales have become very popular with the almost universal introduction of Scholars' Record Cards. Some of these rating scales have been far too ambitious, and have attempted to rate qualities which may not really exist as psychological entities. Moreover, the fluidity of the young child's nature often means that a judgment of his "character" is merely a judgment on how he reacts in a particular lesson, or to a particular teacher, or at a particular time or period. Where it is important to arrive at as fair a judgment as possible, it should be remembered that the combining of the judgments of several teachers, independently made, gives a more reliable estimate than that of a single one.

Finally, the value of graphology, mentioned earlier in this chapter in connection with tests of handwriting, as an aid to the diagnosis of character and personality, should not be forgotten.



## CHAPTER IV

### CHOOSING TESTS

If a test is chosen and then turns out to be unsuitable, it will involve a waste of time and effort both for children and teacher, and if the test is one which has to be purchased from the publishers, especially if it happens to be an intelligence test (as these are expensive), there is also a waste of money. Moreover, where test material has to be ordered on the "Requisition", a test programme has to be organized well ahead, and a failure to select wisely may frustrate a scheme and delay its satisfactory execution for a whole school year.

#### **Priority for Intelligence or Attainment Tests.**

Since attainment depends on capacity, it would seem at first sight that intelligence tests should precede attainment tests, for the level of attainment actually achieved should be related to the innate capacity as revealed by the intelligence test.

Many attainment tests, however, do not require any expense beyond the duplication of test blanks,<sup>1</sup> which can be done in school, or the acquisition of a volume on mental testing. Others can be obtained from publishers at rates much lower than those for intelligence tests. Therefore, since the attainment level will at least give some guide to the intelligence level (and in most efficiently run schools will not diverge from it a great deal), if attainment tests are given first, these will provide some idea of the mental age range likely to be involved.

In any case, a teacher who has not had much experience in using standardized tests should begin by using attainment

<sup>1</sup> Care should be taken to see that such duplication does not infringe copyright.

tests in order to familiarize himself with the principles involved in such testing. It is often a surprise to new users of standardized tests to discover how great a range of attainment can exist in one class. Thus a small junior department with a single form entry in an "average" neighbourhood would normally have in the first class a mental age range of from 4 years at the lower end, rising to 6 years or more at the top. This is illustrated in Figure I on page 70. Where in addition there are children with specific disabilities in certain subjects, whose attainments do not keep step with mental capacity, the attainments age spread is probably larger. This difficulty, for such it is bound to be in class teaching, is modified somewhat where the school is larger and classes are "streamed".

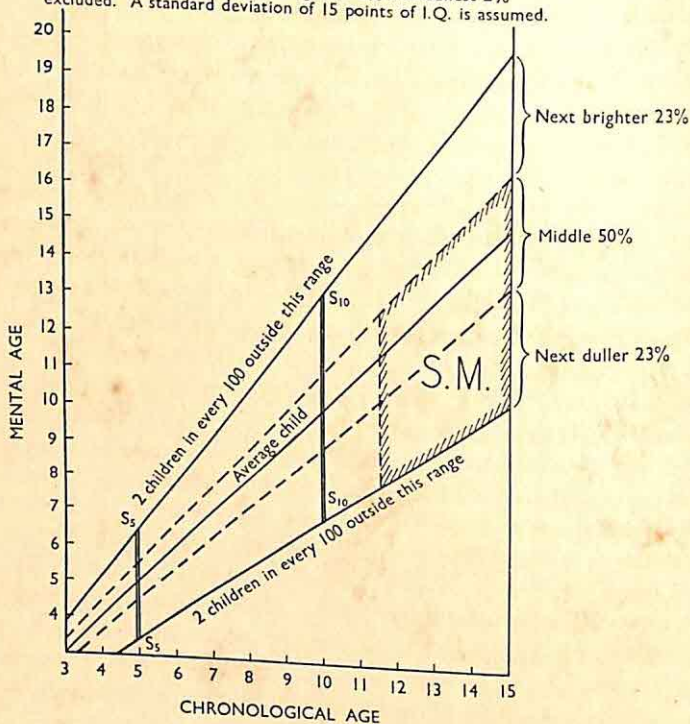
The Secondary Modern School has an equally difficult problem with regard to mental age, for although the brighter children may have left for other schools, what might be called the inertia of the intelligence quotient causes the gap between the mental ages of the dullest and brightest to continue to increase. With reference to Figure I it is assumed that the dullest 2% have to be catered for in Special Schools. This is a very optimistic assumption, for few authorities have made as much special provision as this, and many have made none at all, so that the dullest children, who may have mental ages running down to 8 years, usually have to be taken into consideration. Further, very few authorities make provision for all the brightest 25%, so that the mental age range is also extended at the upper end. If it is realized in addition that not all schools fit into this normal distribution, it is easily seen that the selection of an appropriate intelligence test is a task which requires a great deal of thought.

### **Transfer or Common Entrance Results.**

There is a growing practice for authorities to examine all the children in their areas at 11+ and to make the results available to the schools concerned. Where this is done the Secondary Schools have valuable data for grading purposes, though where



Graph showing the increasing mental age range with increasing chronological age. Brightest 2% and dullest 2% excluded. A standard deviation of 15 points of I.Q. is assumed.



- $S_5$   $S_5$  - Spread of mental ages at 5 years  
 $S_{10}$   $S_{10}$  - Spread of mental ages at 10 years  
 S M Proportion of children likely to be left in the Secondary Modern School. (This assumes that the brightest 25% proceed to Grammar or Technical Schools.)

Figure I

verbal tests have been used the results for the duller children are not likely to be very discriminating. The Junior Departments, of course, receive test results only for children who are about to leave them. These results, however, can be used for deductive purposes, especially if successive years yield fairly consistent distributions. Suppose for instance that the mean

I.Q. is approximately 90 for a certain school over a number of years, then one could reasonably assume that the mean I.Q. for each preceding year is likely to be 90. The range of I.Q.'s can also be examined and the probable range for earlier years thus determined. Suppose it is 70 to 110, then since  $I.Q. = \frac{M.A.}{C.A.} \times 100$  if we want the relevant data for 7 years we have  $M.A. = I.Q. \times 7 \div 100$ , from which we deduce a mental age range of from 4.9 years to 7.7 years with a mean of 6.3 years. Similar data could be worked out for other classes. The chronological age would, of course, be the mean chronological age of the class. These deductions would be quite valuable in selecting a suitable test. Moreover, the information would be useful for equating with attainment test results, if it were not practicable to use an intelligence test. Whether such auxiliary information is available or not, it would in most circumstances be better to proceed first with some attainment tests, for the reasons already mentioned, and also because the prior administration of such tests often reveals the desirability of, and frequently the necessity for, an intelligence test.

If it is decided to proceed with attainment tests, the choice of tests must depend on what the teacher considers to be the aim of the testing programme. It is always wiser to make this as simple as possible at first. When a test has been selected it is best to use it consistently, unless, of course, it should prove unsuitable in some way. If this is done it is possible to compare fairly one child with another, an A class with a B class, or one year with another, such as the 6-year-olds with the 7-year-olds. The use of different tests, even though each does provide an attainment age, does not give the fairest comparison. For example, the standardization may vary from one test to another, or even within one test, as was shown by Vernon's restandardization of Burt's test for Scottish children; or the test might not provide a sample of quite the same aspect of the work, as when one test of word recognition consists mainly of "look and say" words and another mainly of "phonic" units.



## Reading Tests.

Let us first consider reading. The aim of reading is of course to comprehend the written word, but this aim is not achieved immediately. Even though a child comprehends as he learns to read it is some time before he has a sufficient vocabulary to be *tested* in comprehension, so we find that the fundamental reading test in the Primary School is the graded word reading test. Though some of these tests reach adult level at their upper limits, their main value is for children up to 8 or 9 years, or more precisely for children with reading ages up to 8 or 9 years, for backward children have to be tested with material suited to the average younger child. For children beyond this level of achievement the comprehension test becomes suitable. There is no difficulty about age range with the word reading test, as it begins with a zero attainment and goes as far as the Primary School is likely to find necessary. Comprehension tests, however, will be found to have limited age ranges, for example, 7-11 years, 8-12 years, etc.

*The question of age range raises an important point which is applicable to all tests, including intelligence tests, with a circumscribed age range. The test is usually most effective at the middle of the range and, unless it has been specially constructed, is least effective at the ends. One should always beware of placing much reliance on a result coming close to either end of the scale, for if the child scores only one or two points it may be too difficult for him, and he should have an easier test, or if he scores nearly all the points possible, the test may be too easy, and he should have a more difficult test to give him more scope. In this connection I recall that a teacher once condemned "testing" altogether because he had given an intelligence test to a child and found a certain high intelligence quotient; he had given the same test three or four years later and found the intelligence quotient had fallen to an average figure. Inquiry showed that on the second occasion the child had "hit the top" of the test, and could not score more because there were*

no more difficult items. His score had not been able to rise to keep pace with his age, and the intelligence quotient was therefore spuriously low. He should have been given a different test with a higher "ceiling".

The reading comprehension test is usually in booklet form and answers are written on paper provided by the teacher. Once the test papers have been bought they can be used many times, so that if the test should prove unsuitable for some children, nothing more is wasted than the writing paper and the quarter of an hour spent on doing the test.

If such a test is too difficult, recourse may be had to the word recognition test, which, though an individual test, is very quickly administered. Should it be deemed important to measure comprehension before the child reaches the maturity required for group testing, a test such as Schonell's Simple Prose Reading ("My Dog") test is useful. This is an individual test which measures comprehension and also speed and accuracy of reading. Naturally it takes more time if all these measures are to be obtained at once, but the full test is more diagnostic.

### Other Language Tests.

Since reading is a fundamental accomplishment, it should obviously take precedence over English tests. The result given by a reading test should also frequently give guidance as to which further English tests would be most profitable, that is, whether the testing should be more concentrated as in a diagnostic testing programme, or, where the results seem satisfactory, whether consideration should be given to the advantages to be gained by going on to more scientific ways of assessing composition, for instance, or the measurement of vocabulary. The availability of the reading age also makes possible the consideration of verbal versus non-verbal intelligence tests, for, as was mentioned in Chapter III, there are doubts about the validity of verbal intelligence tests for children with low reading ages. If the children *can* read, however, a verbal test will



usually differentiate between children, and show which are brighter and which duller, but it may not give an accurate intelligence quotient or mental age.

### Spelling Tests.

Spelling tests are here mentioned separately from reading and other English tests because, in my opinion, they are rarely worth giving unless they are to be used in connection with a scheme of teaching spelling such as Schonell's<sup>1</sup> or Boyd's<sup>1</sup>. In the reaction against mechanical teaching earlier in the century many adopted the policy of teaching spelling "incidentally". This was a move in the right direction, as it turned attention away from bizarre words, which were often completely outside the child's natural vocabulary, towards the words he actually used. The method of teaching spelling "incidentally" has, however, unpredictable results, and is wasteful of teaching effort.

Modern spelling schemes, presenting as they do groups of words of similar visual pattern, groups of words containing similar auditory components, and separating words which are likely to be confused if presented together, make learning to spell a much more economical business. Since they also present words in groups suitable for different ages, determined by analysis of children's written work, the teaching of spelling can be quite scientific, related to the children's needs, and economizing effort on the part of both teacher and taught. Unless, therefore, teachers are willing to adopt the policy of making the teaching of spelling as scientific as possible, there seems little point in giving standardized spelling tests, for all they reveal is the words in the test which the child does or does not know, and to correct the errors merely ruins the test for future use.

It might, of course, be argued that no standardized test is worth giving unless the teacher intends to improve the efficiency of his teaching. In general this is sound, but the point here is that increased effort in spelling is unlikely to give

<sup>1</sup> See Bibliography.

a good return for energy expended unless based on a definite scheme. Moreover, a perusal of a page or two of the introduction of Schonell's book on spelling, which is founded on considerable research, should convince the most optimistic teacher of the impossibility of devising such a scheme himself. In reading and arithmetic there are many good schemes and textbooks, and merely to spend more time with these would invariably produce improvement—though no doubt diagnostic and remedial work would do much more—but in spelling without a scheme energy is easily dissipated.

It has been mentioned in Chapter III that the best spelling test is a dictated one, and this type should always be used unless there are special circumstances such as might arise in a secondary school transfer examination, where, for instance, separate invigilators might dictate with accents unfamiliar to the children, and thus confuse them. Here the underlining of alternatives is perhaps the fairest way, though it probably favours the good visualizers, and does not require the child to *write* the words, which after all is the essence of spelling. The age range problem does not arise in this field, as standardized spelling tests are usually intended for children whose ages range from 5 or 6 to 13 or 14 years.

### Arithmetic Tests.

Here again policy must be considered before any testing is done. If one is anxious to proceed to an intelligence survey, it will probably be best to begin with a test of "problem" arithmetic or a graded mental test. Nevertheless, the mechanical side is fundamental, and unless a child can handle the basic addition and subtraction combinations, the "tables" and the "four rules", at least up to the normal capacity for his age and general ability, he will be severely handicapped. Very young children cannot be tested by a group test owing to "setting down" difficulties. For children of up to about 7 years one is restricted to tests like Ballard's "One Minute" oral addition and subtraction tests, or the lower



end of Burt's "Graded Oral Tests". For older children there is no dearth of suitable tests, but, in choosing a "mechanical" test, teachers should be clear in their own minds whether they want to test knowledge of basic number combinations and proficiency in the four rules (i.e. number only), or whether they wish to include money, capacity, decimals, fractions, etc. The former type would probably be best if nothing were known of the standard of work relative to other schools, or to mental age, whereas the latter would be appropriate if it were already known that the fundamentals were sound. Examination of any of these tests will show that they are not really suitable for arithmetic ages below 7 years. Some tests have been prepared in two comparable forms, e.g. Form A and Form B in the Schonell "Essential" Tests. The purpose of the two forms is to eliminate practice effects where the same test is used repeatedly. This might be of considerable advantage where intensive coaching is taking place, but usually standardized tests are not given so frequently as to cause difficulties of this kind. Apart from the use of alternative forms of the same test where this is really necessary, it is best to keep to one test, so long as it is not obviously unsuitable. As a rule the value of a test increases with our experience in handling it (some reasons for keeping to the same test have already been given).

The purpose of arithmetic is, of course, essentially practical, so that the "problem" type of test should always be given if we want to be able to make a final judgment on the value of our teaching. Many children, who can work out a sum if it is presented in a stereotyped form, are confounded when it is presented in a form which gives little or no clue to the method of attack. It must be remembered, of course, that there is a considerable correlation between "problem" arithmetic and general intelligence, and intelligence must therefore always be considered when making any such judgment on the effectiveness of teaching.

The problem arithmetic test is often arranged like the word recognition test, with so many items for each year. It is not

therefore necessary to give the whole test. Children should begin at what is considered an appropriate year for them, but if that is found to be too difficult they should tackle the earlier items. The starting level is actually the highest year in which they are correct in *all* the items. They are given credit for any earlier items which presumably they would be able to do correctly. The child should also go on until he fails all the items in a particular year, and be given credit for all items correctly answered. There is some danger of fatigue in this type of test, and if this is likely to occur the test should be spread over two lessons on different days.

### Intelligence Tests.

Basic 3-R tests have now been considered. If some of these have been administered, data should then be available to give guidance in choosing a suitable intelligence test. The combined results of a reading<sup>1</sup> and problem arithmetic test should probably give the best indication for this purpose. From these results the mean or average attainment age should be obtained, and also the range of attainment ages. It may be that the range is exceptionally large because of the presence of one or two outstanding children, found either at the top or bottom of the range, or at both ends. It may sometimes be advisable to choose a test that would not be suitable for these children, and to examine them with another test, if by such exclusion the first test would be more suitable for the majority of the class. This should be a test which would place the children well within the mental age range of the test. For example, if a class is known to have a mean attainment age of 7 years, it would be unwise to choose an intelligence test with a mental age range of, say, 7-11 years, as half the children would find the test too difficult. If on the other hand the mean attainment age were  $8\frac{1}{2}$  years, and the range of attainments spread from 7 to 10 years, the

<sup>1</sup> "Mechanical" or "word recognition" tests of reading are found to give good correlations with intelligence, so this type could be used for the above purpose if a "comprehension" test is not feasible.



same test would be more suitable, though there would be a risk of the duller children finding the test too difficult.

An exception to this procedure arises where a test is standardized for children of a certain narrow age-group. The "Moray House Picture Intelligence Test", for instance, is for children of  $6\frac{1}{2}$  to  $8\frac{1}{2}$  years. It should be noted that the scores are applicable only to children with *chronological* ages lying within these limits, but a full range of intelligence quotients is available for children taking this test, though it is slightly less discriminative at the extremes, that is to say, for the duller  $6\frac{1}{2}$ 's and the brightest  $8\frac{1}{2}$ 's. Thus while without doubt a test of this type could be safely given to children between the ages of 7 and 8, a preliminary attainment test giving mean and range would tentatively show whether it would be wise to use it outside the middle range. If the results of the preliminary attainment test suggest that the children are brighter than average, that is to say, that their attainment test ages are higher than their chronological ages, then it would most probably be suitable down to the lower limit of  $6\frac{1}{2}$  but not up to the upper limit of  $8\frac{1}{2}$ . If on the other hand preliminary testing shows the children to be duller than average, the test would be at its best for the middle and upper age range, say from  $7-8\frac{1}{2}$  years.

### Examiner's Booklet.

It is always advisable for test administrators to study a copy of the Examiner's Instructions, which almost invariably contains a marking key and table of norms. If the table of norms awards mental ages (or intelligence quotients) for very low scores such as 1, 2, 3, etc., these lower figures should not be taken too seriously. A child might easily have obtained such a small score entirely by chance. The lower useful limit of the test should be taken at scores of from 5 to 10.

The consideration of a suitable test, then, especially from the point of view of mental age range, should include an examination of the published norms. It is much better, and would be a big saving of time and money, to reject a test at the

initial stages, rather than to spend money on the test itself, only to find it unsuitable for a proportion of the children.

### Verbal or Non-Verbal Tests.

Some comments have already been made on this topic in Chapter III. Dull and backward children are best examined with non-verbal tests if there is any question of reading disability, and it must be remembered that a reading age of  $9\frac{1}{2}$  is very probably a minimal reading age for the satisfactory functioning of a group verbal test. With children who cannot read the non-verbal test is, of course, a necessity, so long as group testing has to be done. The choice between these two tests will therefore depend on the age at which it is desired to test, and the mental calibre of the children. In "duller" schools, and for younger children, the emphasis will be on non-verbal tests. In above average schools where the children are old enough, the verbal test will probably prove satisfactory.

The testing policy, however, must also be considered. If intelligence testing is to be carried out consistently in a school as an organic part of the administration it is probably best to take annual cross-sections and examine, for example, all the 7-year-olds or all the 9-year-olds (or both). Thus in course of time each child in the school has one or more intelligence measurement. On the other hand, it may be desirable to examine every child in the school at approximately the same time. This usually presents something of a problem, because it involves trying to reconcile two conflicting principles. One of these is to keep the mental age range of the test as narrow as possible, so as to discriminate to the highest possible extent among the children, and the other is to use the same test for all the children so that the basic situation will be strictly comparable for all those being tested. This second method would increase the mental age range and work directly against the finer discrimination, for it would probably involve grouping the children together closely at the ends of the scale. That is to say, the bright children would be clearly separated from



the average and the average from the dull, but it might fail to distinguish between the dull and the "educationally sub-normal", or separate the bright from the brilliant.

In practice, therefore, if a whole school or some large part of it has to be examined at the same time, it is better to compromise and use two tests. Usually a verbal test for the brighter and/or older children and a non-verbal test for the younger and/or duller children would be the best choice.

### **Another Method of Selecting an Intelligence Test.**

If there were some doubt as to how such a large survey would work out in practice, it might in the long run be more profitable to make a tentative choice of tests, and sample the effectiveness of the choice. The tests could be given to representative groups of children from each year whose efforts would enable the examiner to assess the relative possibilities of the tests. The same method could, of course, be tried with a narrower age-group, and might in some circumstances be worth while.

### **Individual Tests.**

With older children, for example from the secondary stage onwards, there seems little to be gained by using individual tests, apart, of course, from special ones such as the "performance" type. The group test will almost always give a satisfactory result, unless some very special emotional condition or character disorder is present. Such a condition would almost certainly manifest itself also in the individual test, but in that case its actual operation would be more apparent, and in addition a qualitative assessment of the child could be made.

With younger children, however, the case is different. In the first place time limits are not generally imposed, and the presupposed "test sophistication" and "work maturity" level are not involved in the individual test. The test items are on the whole more nearly related to real life situations, and the examiner is in a position to encourage and stimulate the child

to produce his maximum effort. The non-verbal group test, which is the only alternative for younger children, is sometimes rather abstract in part, and while being able to think in abstract terms is the hallmark of general intelligence, if the abstraction is too remote from the child's experience, it may fail to interest him and therefore give a poor measurement. One bright youngster of 7 who gained an excellent score on a certain group test had a zero score on one of the sub-tests. When asked about this he said he thought it was silly, and had made no effort to understand what it was about. The sub-test was an "analogies" test involving geometrical figures, and to this bright independent child it was merely silly. Many of us would be inclined to approve of his reaction as an intelligent one for his age, yet he was penalized considerably in the intelligence quotient he earned from the test.

There is no doubt that for younger children the individual test is superior, and for very young children it is, of course, a necessity. It has the additional advantage of costing almost nothing to administer, but it does take some time. The Valentine test, for instance, takes on the average twenty minutes for junior school children, and half an hour for infants who have been in school about six months. Many, however, would claim that the time spent is fully justified. There is no doubt that, so far as the youngest children are concerned, such individual testing is the only really satisfactory method. It could be claimed that the possession of such a test, and the ability to administer it, is in any case almost a necessity for checking doubtful test results, and for the preliminary distinguishing of "educationally sub-normal" children at as early an age as possible.

### The Secondary Schools.

So far in this chapter it is mainly the Primary Schools that have been considered. When children are admitted to the *Secondary Grammar School* their intelligence and attainment quotients are usually obtained from the Authorities' Transfer



Examination which they sit at the age of 11+. Some schools grade their new entrants with the help of the data so supplied, and make adjustments at the end of or during the first year. It would not be surprising if the grammar schools do not find this information entirely satisfactory. The Transfer Examination aims at separating the brighter from the less bright, and no doubt it does this with a fair degree of efficiency. But such a test, designed as it is to select the top 15 to 25%, will not at the same time be able to differentiate accurately between the selected children in this top group. This fact, taken in conjunction with character factors (industry, perseverance, etc.), interest, and the possible maturing of special abilities, will show that the usual 11+ intelligence test cannot possibly give the best prognosis of grammar school success *within the grammar school itself*—that is, among the selected children. It is here that the special tests developed by Dr. Earle should be particularly valuable. These have already been mentioned in Chapters I and III. If any grammar school teacher has felt disappointed in the efficacy of the 11+ examination in predicting success *at the grammar school level*, he should reconsider the matter in the light of what has been said above and give Earle's tests a trial. There is no doubt in my mind as to their great value, and they will do for the grammar school what the test of general intelligence will do for the infant and junior school. It is possible that a misunderstanding of the 11+ examination, as well as a general conservatism, has been a deterrent to the use of the standardized test at this stage.

Standardized *attainment* tests for grammar schools have not had much attention in the past, though it would seem that there is a definite field for the use of the "new-type" examination. No doubt the age performance principle is difficult to apply, as the practices of different schools vary so much. Nevertheless, much might be done if the value of such tests were appreciated as an aid to measuring progress and detecting weaknesses and gaps in knowledge due to absences or other causes. One or two examples of attempts to formulate

standardized tests are given in Chapter VIII, which might serve to stimulate the interested teacher, either alone or in co-operation with colleagues in other schools, to build on what has already been done. Doubtless also the study of the graded word reading test, the comprehension test, and the mechanical and problem arithmetic test would provide useful models for the language and mathematics teachers respectively. The latter would in many schools find standardized arithmetic tests and diagnostic arithmetic tests useful as they stand, so far as lower forms are concerned. The example given in Chapter I of the boy who made good in mathematics,<sup>1</sup> after being coached with the aid of diagnostic and standardized tests, should give point to these remarks.

The *Technical School* is to be considerably developed in the future. Here again the teacher must not expect too much from the test of general intelligence. Such a test is, of course, necessary for a preliminary screening, but children with more than the minimal intelligence having been chosen to follow the technical course, special tests are necessary. A great deal of research is going on, and more help will be available in this direction in the future. Chapters III and VIII refer to the material already available, and there is great need to experiment with this. Interested teachers would find it well worth while to consult volumes of the *British Journal of Educational Psychology* from 1946 onwards.

The *Secondary Modern School* has doubtless the greatest spread of ability, as Figure I shows, while the spread of attainment is often still larger. Moreover, conditions for the Secondary Modern children vary enormously, for some of them are still found as seniors in non-reorganized all-age schools, while at the other extreme some are in large modern schools with four or five "streams". Thus the problem of advising on the choice of tests is a complex one. Teachers in smaller schools will no doubt find much of what has been said in relation to Primary Schools applicable to their prob-

<sup>1</sup> Page 12.



lems. On the other hand, teachers in large Secondary Modern Schools will, it is hoped, find what has been said in relation to the Grammar and Technical School of some use, for the division between these schools and the Modern School is arbitrary, and children who in one area might be found in the Technical and Grammar School, may in another be found in the Modern School. The Secondary Modern School, because of the large spread of ability and attainment likely to be found there, has the most difficult organization problem of all schools. Very few of these schools escape having a proportion of virtually illiterate children, and few also find themselves without the bright children who might conceivably be well suited to a grammar or technical education. The school therefore has the complicated task of attempting to cater for all and to provide a suitable education for each *level* of ability, as well as for the various supposed *types* of ability.

No doubt all Secondary Modern Schools would find it profitable to conduct standardized attainment tests, beginning with the 3 R's. For the poor readers the word recognition test will still be serviceable, and at times necessary. It will be found that standardized attainment tests frequently do not have an age range beyond 13 or 14 years. This has to some extent been conditioned in the past by the statutory school leaving age of 14 years. On the other hand, a case can be made for regarding 3-R tests, at least on the mechanical side, as tool subjects which should be properly "elementary" subjects, mastered by the majority of children in the Primary School, or in the early stages of the Secondary, and therefore not needed beyond the age of 13 or so. Reading comprehension, for instance, becomes increasingly dependent on vocabulary, and no doubt a standardized vocabulary test, especially an extensive one like the Watts test, becomes more serviceable at the secondary stage. Incidentally, this test is found to correlate highly with verbal intelligence tests, and can therefore be used as a supplementary estimate of intelligence before an intelligence test is chosen, or as a check to other results.

It is assumed that most secondary schools will have a verbal intelligence quotient available for each child, derived from the Transfer Examination at 11+. For schools where this is not so there are excellent well-standardized tests available, and for those where expense is a principal consideration, some of these are provided in such a form, that answers can be written on separate papers, and the test sheets themselves used over again.<sup>1</sup> It would be *advisable* to retest all the new entrants to the Secondary Modern School with a non-verbal test where the verbal intelligence quotient is below 80, but it would seem to be *vital* to do this where the reading age is, in addition, below 9½ years. Only by this means can we be quite sure that some children are not being regarded as dull and backward, who may be nothing of the sort. Since the general introduction of the 11+ intelligence test I have seen numbers of children who have been supposed to be educationally subnormal, but who have turned out to be anything from merely dull up to average, the supposed deficiency being nothing more than retardation in reading due to emotional blockage, absence, disturbed educational history, or some specific disability which could be revealed by diagnostic testing. A trivial calculation with the intelligence quotient formula for the group of children in the first form, with intelligence quotients from 70-80, indicates that such a non-verbal test should cater for mental ages of approximately 8-9½ years.

With the data available from attainment and intelligence tests in the first year, the organization of the Secondary Modern School should be relatively easy. Help in this respect should be obtained from Chapter VII. During the first two years there is ample opportunity for further study of the development of the children's special abilities. Further testing with performance or other technical tests, or with one of Earl's "Duplex" tests, will help to direct children into appropriate courses. It should not be overlooked that the Grammar and Technical Schools cater for different types of

<sup>1</sup> E.g. "Northern" and other tests prepared by Tomlinson.  
(G 908)



ability at a fairly high level of general intelligence. The same types, with less general ability, presumably exist in the Modern School, and at least in the larger schools could be dealt with in streams appropriate to their special bents and interests. It is the invariable practice in boys' schools, for instance, for all to do some form of craftwork, but many of these boys have no real ability, especially in the more technical aspects of the work, and the associated mechanical drawing. Such pupils might very well be given a modified course of "home repairs" (tap-washers, electric fuses, etc.) or an alternative course. The same thing applies in the case of girls. Some are good at design and pattern drafting, but others, while being excellent needle-women, have no flair for this. Here technical and performance tests would no doubt be of much help, but it is obvious that some considerable thought would need to be given to the matter, or disillusionment might follow, as in the case of the too enthusiastic user of the general intelligence test. There is considerable need for teachers to experiment and make known their findings in the educational journals.

To sum up, the following points should be remembered when choosing tests:

1. Use any available data such as results of secondary selection examination. For younger children, where the school population is stable, the mean and range will probably be similar.
2. Proceed from the simple to the more complex, e.g. reading (word recognition) or reading (comprehension) before diagnostic tests, composition scales, etc.
3. Use attainment tests first, and then estimate probable mean mental age and range.
4. Where possible, achieve this estimate by means of reading comprehension (or word recognition if such only is available) and problem arithmetic tests, rather than by spelling or mechanical arithmetic tests, as the former are usually more "intelligence saturated".
5. From such an estimate use an intelligence test whose

middle range comes as near as possible to the mean of the attainment test results.

6. As far as possible use the intelligence test with the narrowest age range for the number of items in the test. A test with, for example, 75 items for a 2 year age range would, other things being equal, give much finer discrimination than a test with a 7 year age range for the same number of items.
7. Always study the Examiner's Instructions and Norms before ordering the tests in bulk. The former costs only a shilling or so and will show the mental age (or intelligence quotient) range, from which it can be seen where the test scores cease to be discriminative.
8. It might, in certain circumstances, be advantageous to "sample" the children (2 brightest, 2 average, 2 dullest) with one or more tests before making a final choice.
9. Where test results seem to contradict your subjective opinion, give supplementary tests (individual, performance, or another group test) and bring the matter to an issue rather than reject the test.



## CHAPTER V

### DIAGNOSTIC TESTS

In one sense all standardized tests are diagnostic, for the word diagnosis is defined as "accurate investigation and discrimination by means of data", but what are usually called diagnostic tests are specially designed tests of a more refined nature, aimed at pin-pointing a disability either quantitatively or qualitatively, or both. Thus the ordinary standardized arithmetic test might indicate that a certain child is backward in arithmetic, and it should also indicate by how much he is backward compared with the average child of his own age. The diagnostic test, on the other hand, would aim at an *analysis* of the backwardness and might reveal, as is often the case, that the child has an imperfect knowledge of addition and subtraction combinations. It would also differentiate between the child who is merely slow, and probably has to count on his fingers to build up the basic combinations, and the child who is not slow but consistently makes the same errors, such as calling the sum of 8 and 9 always 15. Another child may have difficulty with division, possibly because there is some underlying difficulty in the process of subtraction which he has never mastered. The diagnostic test would show that the prime weakness lies in the subtraction, which must be corrected before the division can profitably be tackled.

#### **Prior Need of Standardized Tests.**

In general the use of standardized intelligence and attainment tests precedes the use of diagnostic tests. It would seem that, except in special circumstances, if a child has been given

intelligence and attainment tests, and the attainment age more or less matches the mental age, then however backward he is, he is doing as well as can be expected and diagnostic tests are not required. No doubt such tests would often reveal weaknesses which could be remedied, but since I have in mind the ordinary busy teacher who has to cope with the needs of all the children in his class, it will seldom be practical to deal with any but the worst misfits, whose disability is a serious handicap in keeping up with their own potential. It would be a serious misfortune if the diagnostic test played into the hands of the obsessional or perfectionist. A ready acceptance of one's own inferiorities and those of others is a sign of mental health, and to accept the fact that a certain child is a "duffer", and to think no worse of him for it, is necessary and just. On the other hand, it is unjust to label him thus if his disability is one which arises from some reparable defect, and it is here that the standardized intelligence and attainment tests have their value in helping to form a sound judgment. After comparing capacity and achievement, we can then decide whether the diagnostic test is likely to be a useful aid or merely an unnecessary refinement.

### Diagnostic Possibilities of Ordinary Standardized Tests.

It has been mentioned in other chapters that the ordinary standardized test can be used for more than the mere numerical result which it gives, and it now seems appropriate to enlarge a little on this point. Arithmetic tests of the four fundamental rules, for example, may reveal a lack of uniformity in progress. There may be a greater weakness in addition than in multiplication. Judging from my own experience I should say that this seems to be quite a common state of affairs, probably a result of the Junior Department's concentration on "the tables", on the assumption that the basic addition and subtraction combinations have all been mastered in the Infant Department.



With separately scored tests for each "rule" it can be clearly shown whether any such process of concentration is unconsciously taking place in the child's mind, and whether the weakness is a general tendency or merely a difficulty of certain children.

Sometimes subtraction is the weakness. Such a test shows at once the children who are confused about "borrowing" or those who have difficulty with "0", for instance. In "long" multiplication again, it may be obvious that a child is confused as to whether to begin to multiply from the right or from the left. In division there may be errors in "carrying", or a child may be found to be consistently leaving remainders which are larger than the divisor. Scrutiny of the children's tests along these lines at once becomes diagnostic. Whether the papers for a whole class should be so examined, or those of certain children only, can be decided from the relationship between the mean arithmetic age and the mean mental age of the class. If these two are more or less equal it can be assumed that the level of the class as a whole is satisfactory, and attention can then be paid to the papers of those individuals whose scores lag considerably behind what would be expected from their mental ages. If the whole class is retarded, a search for common types of error should be most profitable, as they can be dealt with as a class problem.

It cannot be urged too strongly that the qualitative aspect of standardized tests should always be kept in mind, and at the same time their power (as with the results from tests of the four rules) to make a differential diagnosis should be noted. An interesting example of this arises in the case of Watts's Vocabulary Tests. There are five tests, consisting of Tests 1 and 2, each having 100 common names, Test 3, 100 common verbs, and Tests 4 and 5, each containing 100 common adjectives. Watts found that the poorly taught child, especially when in a slum school, did relatively much worse at the test consisting of common verbs. His conclusion seems to be that cultural background makes a difference here, and that the

child who reads and writes early, and who is in contact with more literate people, uses more verbs and relies less on gesture. Such a test, useful only for secondary school children, which is able to reveal objectively such interesting differences, is of the greatest value in pointing the way for certain schools to compensate for environmental deficiencies.

Just as a good deal can be obtained from the straightforward mechanical arithmetic test, so the "mechanical" or "word" reading test can provide a mine of diagnostic information, if the *kind* of errors is taken into account after the *number* of errors has been reckoned with to arrive at the reading age. For instance, some children make a consistently faulty directional attack on words, producing "was" as "saw", "on" as "no", "day" as "bay" (letter reversal). One child of my acquaintance had a chronic disability of this kind. She was intelligent, and soon abandoned the simpler of these crude reversals, but persisted in reversing combinations of words. She would read "marigold" as "gold-mary", and, by beginning at the right-hand instead of the left-hand end of the line, would use the words in the line and recombine them to make up a new sentence. She was, of course, too intelligent to read the words from right to left consecutively, as this would have made nonsense, and the trouble only occurred intermittently. A problem of this type should always be considered in relation to "handedness" and also to "eyedness", though left-handedness, or crossed laterality (left eye/right hand or vice versa), will often fail to explain these difficulties. Very often it is found that a child is unaware of his mistaken directional attack, and to make him conscious of this and to give special attention to it results in a considerable and sometimes rapid improvement.

Other children make different kinds of visual errors, reading "cry" for "carry", "shape" for "sharp", etc. Those in this group are often called careless readers, as they seize on predominant visual clues, frequently the initial letters, and then make their own words. In my opinion many of these children



are of the intuitive and impressionable type, who pick up "atmosphere" very readily. This is a valuable function, though it is constantly discounted in school. (An intuitive person can pay a visit and leave quite unaware of much that has been in a room, but full of knowledge acquired by way of the unconscious mind. The good "sensation" type of person would, on the other hand, be able to catalogue the contents of the room.) Intuition is, of course, a drawback in reading where a fairly accurate perception of the pattern of words is needed. I have found that a sympathetic understanding of this difficulty does much to help. These children are often baffled by the number of errors they make, and frequently are emotionally disturbed about their reading. To rebuke and even punish them for using what is for them a natural way of approach is hard and even injurious. To be shown that they have to control this natural inclination to work too quickly from predominant visual clues, which tends to make them read the word "three" as "there", gives them support. Such children, too, may find difficulty in recognizing people, as well as words. For example, they may think they know a person they see in the street when in actual fact the person just resembles someone they have seen at the cinema a day or two earlier. I mention this, not to put forward an all-embracing explanation of visual errors in reading, but to emphasize the fact that what may seem stupidity on a child's part is often a disability for which the child is not responsible.

Analysis of a child's reading errors will sometimes reveal weak auditory analysis, such as the reproduction of "tin" by "ten", "that" by "tat", "sad" by "sed". A parallel examination of spelling errors is worth while, for the child who makes these mistakes is likely to write "cot" for "got" and "comi" for "coming", while the child with visual difficulties is likely to write "lessn" for "lesson", "creul" for "cruel", or produce reversals like "ot" for "to" or "tadl" for "table".

Children often have a good phonic approach but fail to

realize that there are many words which have to be learned by "look and say". This deficiency often betrays itself in their attempts to build up words by sound which unfortunately do not yield to such a method.

High-frequency deafness will tend to reveal itself in speech defects, and confusion and omission of such sounds as "d", "t", "sh", "th", "v", and "f" both in reading and spelling.

### **Specially Designed Diagnostic Tests.**

While a great deal of information, in itself adequate to initiate a suitable remedial programme, can often be derived from a careful analysis of the child's efforts on the ordinary standardized test, this is not always so. Since the standardized test is not designed to be specially diagnostic, there may be many difficulties which do not come to light. On the other hand, many difficulties may be revealed, but one may have no idea of their relative significance, so that it is not clear how one should proceed or on which difficulty one should concentrate first. So it is seen that there is a need for additional diagnostic tests. Let us now consider them individually.

#### *Arithmetic.*

The method of diagnosing specific difficulties in arithmetic is somewhat different from that used to diagnose difficulties in reading or spelling. This is due to the fundamental difference between the processes employed in arithmetic and those employed in reading or spelling. While in the latter subjects a vocabulary can be built up by following different methods, in arithmetic the steps that have to be taken are almost invariably the same. Moreover, if the child misses something from absence or inattention, his progress in arithmetic is often fatally impeded, whereas in reading or writing it may not be halted. In arithmetic all progress depends on the child's acquiring a knowledge of the basic number combina-



tions. The ability to do addition and multiplication, for example, depends on a knowledge of the 100 basic addition combinations; the ability to do subtraction and division depends on a knowledge of the 100 basic subtraction combinations. The diagnostic test, therefore, has to begin with an analysis of the child's accuracy in handling these basic combinations. If a child believes that 9 and 5 make 13, every time he adds them together the result is wrong. The value of a diagnostic test lies in the fact that it can go right to the root of the trouble. Once it has succeeded in locating the recurring error all that remains is for the child to correct it and practise calculations involving his difficulty until the error occurs no more. It should be obvious that this method results in a much more efficient training than the method by which the child has to correct a sum by working over a number of other combinations which he knows perfectly well. Moreover, the single correction may not be enough to show the child that he has a tendency to make that particular error, so that when he meets the same combination again he may make the same mistake again.

The two serviceable diagnostic tests are the "Beacon Arithmetic Diagnostic Tests" devised by Dr. Fleming, and the "Schonell Diagnostic Arithmetic Tests" devised by Professor Schonell. The former are qualitative only, and are designed for use with the *Beacon Arithmetic* in connection with the Teachers' Manual. The *Schonell* tests are universally applicable, and have the added advantage of being standardized for children from 7 to 14 years. A brief description of these tests might be desirable and useful as they have such a wide field of application, from the upper classes in the infant department to the lower forms in the secondary school. Test I consists of the 100 basic addition combinations set out in rows and columns so that the child has to write in the answer only. Each combination is set out in both its forms, for example,  $7 + 6$  and  $6 + 7$ , though not, of course, adjacent to each other. The test is also timed, and average times taken by children of each age between 8 and 13 years are given. This is ex-

tremely valuable, for the test thus provides evidence not only of the accuracy of the learning of the number bonds, but also of any inaccuracy or faulty learning if this has occurred (apart from chance errors), for if the same wrong answer is given to both forms of the combination, it is strong presumptive evidence that the child will consistently reproduce the same error. This can easily be checked by oral questioning. Further, though a child may be able to reproduce the combinations accurately, he may be very slow. This may be evidence of poor capacity to synthesize quickly, perhaps even of mental or finger counting. The ideal is to make the reproduction of the result of the combination quite automatic, and the time taken for the test is an indication of the facility with which this has been achieved. Thus at 8 years the average time for completion of the 100 basic addition facts given by Schonell is 8 minutes, a time which is reduced to 3 minutes by the age of 13. It may be of interest to record that many intelligent adults, on taking an interest in these tests, find that their reproduction of some combinations is not immediate; a certain negative feeling tone delays the prompt reproduction, or the intermediate process of building to 10 has to be done before the final result is achieved. A few minutes' attention to these unconscious weaknesses finally removes them as a rule.

Schonell's tests 2, 3, and 4 deal similarly with basic subtraction, multiplication, and division combinations respectively. Children rarely make even progress in the four processes, and a tabulation of scores will usually reveal children who need extra practice in one process to bring them up to a fair all-round level. Research has shown that certain combinations prove more difficult to master than others, and Schonell has therefore provided a miscellaneous test (Test 5), consisting of the most difficult items in the four processes. In this test the different processes are mixed so that the child has to add, subtract, multiply, and divide in random order. It is found that some children, who have not become sufficiently conscious of the significance of the four processes, do not know when to



use which process. These five tests, then, are extraordinarily valuable, and provide accurate measures of the child's competence in the foundations of all arithmetical skill.

To these basic tests Professor Schonell has added three supplementary tests (X, Y, and Z), an examination of which excites one's admiration at the thoroughness of the diagnostic procedure which this author has devised. Supplementary Test X is what is described as a test of higher decade addition. It is observed that some children who know the simple basic combinations are uncertain of them when they appear in compounds; thus the child who knows the answer to  $6 + 9$  may be unable to add 36 and 9. In Test X Schonell has selected 100 of the most difficult of these "higher decade combinations" (basing his choice on the frequency of errors), so that it is often possible to discover if these errors occur in column addition. Supplementary Test Y is similarly arranged to test the addition combinations involved in "carrying" in simple multiplication. Thus to multiply 8865 by 6 the child says  $6 \times 5 = 30$ , put down 0 and carry 3;  $6 \times 6 = 36$ , 36 and 3 are 39, put down 9 and carry 3;  $6 \times 8 = 48$ , 48 and 3 are 51, put down 1 and carry 5;  $6 \times 8 = 48$ , 48 and 5 are 53. This test consists of the 130 most important of these combinations. Supplementary Test Z consists of 110 of the most simple division facts with remainders. All these supplementary tests can be given orally.

Tests 6 to 12 consist of graded addition, subtraction, multiplication and division, long division (easy steps and harder steps), and mental arithmetic. A description of Test 6, Graded Addition, will give an idea of the comprehensive nature of the remainder of these tests. There are fourteen "steps" in the test, with four examples of the same type in each step, these being graded in order of difficulty. Examples of the first four steps are as follows:

14	10	31	543
<u>3</u>	<u>15</u>	<u>66</u>	<u>126</u>
—	—	—	—

It will be seen that none of these examples involves carrying, and the progression is from units in one line to hundreds in both lines. The second step involves 0's in each example. From the fifth to the ninth step these types are repeated, but carrying is introduced once, first in the units place, then in the hundreds place. Step ten has two carrying figures and step twelve has three. The eleventh, thirteenth, and fourteenth steps contain examples of progressively more difficult column addition, evolving from examples of the first to examples of the second of the types shown below:

74	951
56	382
43	467
<hr/>	539
—	<hr/>
	196
	<hr/>

This test and the others in the series are invaluable for analysing the work of backward children, for the exact stage which they have reached, or the type of error they are inclined to make, is at once apparent.

Additional valuable features of the tests are that they are provided with norms, both for time and accuracy, for ages from 7 to 14 years, and a "weakness" figure is given. The norms are of course *average* achievements; some children will do better and some worse than the average; the "weakness" figure, provided for both speed and accuracy, is a danger level indicating which children are in serious need of special help. The teacher of arithmetic who has not yet met these excellent tests, or who has failed to recognize their significance, is urged to discover for himself their simple character and lucid presentation.

### *Spelling.*

Research has shown that the school has a greater influence on spelling than on any other subject. That is to say, an



improvement in teaching produces a greater improvement in spelling than in any other basic subject. The value of the standardized spelling test in connection with a sound scheme of teaching spelling has been emphasized in Chapter IV. Such schemes are, of course, devised for the average child, and the backward speller needs some additional help. A diagnostic programme, designed to detect different types of spelling disability, enables us to arrange this additional help intelligently and economically.

Some hint has already been given that the ordinary spelling test can yield valuable diagnostic evidence. In many cases, however, it would not yield sufficient unequivocal evidence on which to base remedial work, and it is in these more difficult cases that the more specific diagnostic test has its place. For simple workable diagnostic tests we are again indebted to Professor Schonell. There are three tests dealt with fully in *Backwardness in the Basic Subjects*, namely Immediate Recall of Three-letter Words (Visual Presentation), Immediate Recall of Three-letter Words (Auditory Presentation), and Immediate Recall of Nonsense Syllables (Visual Presentation). Three-letter words are arranged in columns of different length and exposed under standardized conditions. The tests are standardized for both boys and girls from 7 to 13 years, and from a comparison of relative achievements on these tests it is often possible to detect a decided disability in the auditory compared with the visual presentation, or vice versa. The spellings also give qualitative evidence of type of error.

Schonell also recommends the use of his Graded Spelling Test (I B Regular Words) and his Graded Dictation Tests, as these respectively provide information about the child's capacity to handle phonic analysis and synthesis, and to deal with continuous material. If these are used following on the ordinary type of standardized spelling test, a schedule of errors may be compiled which would indicate the type of disability from which the child suffers, without having recourse to those special diagnostic tests. Whatever degree of refinement is

undertaken, however, one is faced eventually with the question of how to remedy the defect when it is found, or at least how to help the child to counteract what may be a natural disability. Advice about this is to be found in *Backwardness in the Basic Subjects*, which contains a chapter on teaching backward spellers. Indeed, one might say that no one concerned with the teaching of spelling can afford to neglect Schonell's four chapters dealing with this subject.

### *Reading.*

As has already been mentioned, one of the ordinary word recognition tests, used in a qualitative manner, is the first line of attack in this field. The reader must not overlook the fact that this book does not aim at a complete description of disabilities in the various subjects. This information will be found elsewhere. The main concern here is with tests, and such disabilities as are mentioned are considered only when they are necessary to a study of the available tests. Experience shows that it is very necessary to emphasize this point, as the tests themselves are no substitute for an understanding of the mental processes they aim at analysing. Assuming, then, that my reader has a grasp of the different aspects of mental processes involved in reading, I would suggest that the following tests for weakness in word recognition will be found very illuminating when used with certain types of backward readers.

*Schonell Test R5* is a test of the power of analysis and synthesis of words containing common phonic units. It consists of 80 "regular" words, which have been selected to give a good sampling of short vowels (*van, win*), diphthongs (*ground*), consonantal and vowel digraphs (*bright, road*), and numerous phonic units (*understood, wonderful*).

*Schonell Test R6* is a test of directional attack on words, and consists of forty-eight short words such as would readily indicate tendencies to reading reversals either of whole words,



parts of words, or single letters ("but" for "tub", "gril" for "girl", "big" for "dig").

*Schonell Test R7* is a visual word discrimination test. This consists of twenty-five words selected because they are likely to cause errors. Each word is printed correctly along with five incorrect forms thus:

thred   thraed   threab   threed   thread   threard.

The child is shown the correct word printed on a separate card, and after five seconds' study has to turn to the six given forms and indicate the one he has been shown. This test is useful for confirming a weakness which may already have been suspected from ordinary standardized tests.

In diagnosing reading difficulties, it should be remembered that an analysis of spelling errors may provide corroborative evidence, or give further insight into difficulties or faulty attitudes.

The fact that a child has no difficulties in word recognition does not necessarily mean that he is a good reader. The factors of speed and comprehension have to be considered. The silent reading test usually measures these factors simultaneously, but the Schonell Silent Reading Test (A) has the special merit of dealing with them separately as well as together. The norms give scores for the test when a time of nine minutes is allowed; that is, it measures in a given time achievement which is a composite measure of speed and accuracy. Arrangements are made in this test, however, for the child to mark off the amount done in the given nine minutes, and then to go on and finish the test. The time taken is noted and the number of items correct is a measure of accuracy (the norms are available from 7-11 years). It is thus possible to compare accuracy (comprehension) with speed. Where a silent reading test is not suitable, a test like the Schonell "My Dog" test may be used, which gives a measure of speed and comprehension for oral reading. With this test it is also possible to compare the child's word recognition skill at the same time. The publishers of the Beacon

Infant Readers have an interesting series of "diagnostic" tests, covering word and phrase recognition devised solely on a progress or attainment basis, which are intimately related to this series of readers, so that they indicate when the child may be considered ready to go on to the next stage.

Before leaving the topic of diagnostic reading tests, it should be noted that the process of diagnosis may have to be completed by sensory tests, such as the audiometer test for high-frequency deafness, and the compilation of an educational history for the child. There are excellent schemes for making case records in Burt's *Backward Child* and Schonell's *Backwardness in the Basic Subjects*. The latter offers valuable suggestions concerning exercises for backward readers both in oral and silent work.

### *Written English.*

After difficulties in basic reading and spelling have been overcome, the English teacher's next concern, in the matter of tests, is with written English. Reference has already been made in Chapter III to the application of modern examination procedure in the marking of composition. Here we are concerned with the errors made in the more mechanical aspects of written English. The Schonell Diagnostic English Tests (by F. E. Schonell) are based chiefly on error frequencies. Thus the first test (English Usage) contains fifty-four items covering the commonest types of grammatical mistake, e.g.

5. Please mother . . . I have a pear?

18. My brother and . . . are the only children in our family.

The items are graded in order of difficulty in each test and sufficient examples are given of each type of error to make the tests diagnostically reliable. (The validity and reliability coefficients of each of the tests are given in the Examiner's Handbook; these are very satisfactory for this type of test.)

(G 906)



The second of these tests deals with capital letters and punctuation and includes thirty-six items of the types:

4. Have you seen smith
13. I wish Alice would come said mrs. Jones.

Test 3 is a vocabulary test. It consists of sixty words (nouns, verbs, adjectives, and adverbs) each arranged on the multiple choice principle:

1. A ROOF is the top of . . . a house, a chimney, a tree,  
a cupboard.
25. COURAGEOUS means . . . brave, strong, sturdy,  
stupid.

Sentence structure is the subject of the fourth test. Simple sentences have to be joined and the various joinings are given different values in the marking scale. There are seven groups of simple sentences to be joined, increasing from two in group 1 to four in group 7. The test gives a measure both of the accuracy and variety of connection at the child's command.

If these four tests are to be completely diagnostic they should be allowed unlimited time. Time limits, however, are given, and if they are observed, the tests may be used as ordinary standardized tests. Norms are given from 8 to 14, but the tests are probably suitable only for the brighter pupils of the lower ages given. The tests are useful in secondary schools, and with backward students even beyond the 16 years mentioned as the upper age limit. It is claimed that they have been proved especially useful in grammar schools with pupils who are weak in the mechanics of written English. There is space on the back of the test booklet for a composition on "Home", for which median samples and guidance in marking are given in *Backwardness in the Basic Subjects*.

The "Dingwall" Test in English Usage covers some of the ground of the Schonell Diagnostic Tests. The test is used in two parts, to be given on successive days, and in two forms, one for ages 9-10 and the other for ages 10+-12. The

form for the older children is self-administering, while that for the younger needs a black-board demonstration. Tentative norms only are given, one set for each of the two age-groups. As only small increases in the scores for the sub-tests are shown for the two groups (four of the nine sub-tests show no increase in score from 9-12 years) their value as standardized tests is not great, though, of course, they have a diagnostic value.



## CHAPTER VI

### DISTRIBUTION OF ABILITY

The idea that "if only some appropriate treatment could be found for the dull child he could be made bright" dies very hard. It is perfectly true that here and there a child appears dull because of a neurotic disturbance, and, more frequently, children fail to work up to their potential because of specific disabilities. For these children a diagnostic procedure and appropriate remedial teaching or Child Guidance treatment are needed. Frequently the results of these measures are very good. But for the majority of dull children nothing can be done to make them brighter. Perhaps it would be wise to state explicitly that in this discussion the words "dull" and "bright" are used to refer solely to degrees of intelligence and not to temperamental qualities. Because we live in an age—the so-called scientific age—which owes its success to a pronounced differentiation of the intellectual function, and because so many people have unconsciously come to over-value the intellect, the acceptance of this proposition is difficult for not a few teachers, and where it is accepted it is often regarded as something of a tragedy.

During the last war a soldier fell sick after only a short period of service in the army. He spent a considerable time in hospital, and finally was treated by a psychologist who discovered that he was a high-grade mental defective. His illness was of a hysterical type, induced by the complexities of modern army life, which made very apparent the gap between himself and the average man. In civilian life he ran a firewood business quite successfully. In the army, asked to take down and

reassemble a machine-gun alongside men who seemed to find no particular difficulty in the task, he was bewildered and unnerved. Later on, when an intelligence test became part of the routine procedure when men were called up, cases of this kind did not occur as mental defectives were rejected at the outset. It was the highly mechanized state of the modern fighting force which caused the intelligence test to come into its own, and while the test procedure might have given only a rough grading, it was able to reduce rapidly the very considerable number of misfits.

In spite of the proved practical value of the tests they are still not accepted by many people, among them many who have to undertake them. This is largely owing to ignorance of the aim of the tests. This suspicion is still widespread among teachers, for whereas the life-and-death nature of the struggle made efficiency in the forces of paramount importance in the late war, and therefore forced the authorities to take note of what intelligence tests could do, there is no comparable urgency in schools. Moreover, in a profession where qualities of sympathy and compassion are necessary virtues, there is every opportunity for sentiment to work against efficiency. This is often a good thing, for schools should not ape mass-production factories.

Nevertheless, there is a sphere in which efficiency has a place. A certain kindly and conscientious headmistress of an infant school used to take a group of backward readers in her room for coaching. One of these was a little boy of  $5\frac{3}{4}$ . I found that he had an intelligence quotient of 78 and a mental age of  $4\frac{1}{2}$ . I therefore suggested that his dullness would make it very unprofitable to try to teach him for some considerable time—he would not in fact be at the reading readiness stage until he was almost ready to leave her for the Junior School. She was profoundly shocked at this suggestion and said it was contrary to her principles. She did not like children to leave her school unable to read, and found that if only she gave the time to them, they sooner or later made progress. I pointed out to her



that while here and there one finds a "late bloomer" such an argument is in general false, and whatever progress children make is due to a steady maturing of existing intelligence, each progressing according to his natural endowment, the best indication of which is the intelligence quotient or similar measure of intelligence. The idea of the sudden flourishing of latent intellectual power is not usually borne out by experience. We parted good friends, I sorrowing at my lack of power to convince her, and she to resume her coaching, optimistic of eventual success.

Two years later we met again, and she reminded me of her little pupil and said he had now begun to make progress. I was not surprised, and was wondering how I could tactfully imply "I told you so", without putting it into so many words, only to find her telling me that *she* was right after all! Yet she had spent two years doing what could certainly have been done in six months.

In cases such as these a teacher's charity is misapplied. It is right to be concerned about the child's progress; it is also on the whole a healthy attitude to assume that steady application of effort will bring success. But this alone is not enough. One needs also a fairly close adaptation to reality. Some teachers are too patient, not necessarily with their children, but with conditions they set for themselves. A teacher who puts two years' hard work into the achievement of six months' progress ought to ask himself why this should be. On the other hand, the less patient teacher often becomes cynical. Too much frustration in the attempt to teach a child leads this type of teacher to conclude that the child is unteachable. The one teacher, then, is less than fair to himself, and the other is less than fair to the child. Between these two extremes we have the more sensitive teacher who is not able to identify himself with either of these points of view, and therefore is left in a position of conflict, perhaps moving from the one attitude to the other, or having a good deal of doubt as to his competence or suitability for the work.

All this, now, must be considered as a preamble to the main subject of this chapter, which is the distribution of ability throughout the school population. It would be of inestimable value to all teachers, especially the young ones working with dull children, to realize once for all that intelligence is largely an inherited characteristic; that all children become more intelligent as they grow older up to a maximum at about 15 years; that this maximum is indicated by the intelligence quotient; that the *rate* of growth is indicated by the intelligence quotient; and finally that there are relatively few extremely dull and extremely bright children, relatively more not-so-dull and not-so-bright, with the majority of children coming into the category which is usually called "average". Moreover, once this is understood, and the position of the sample of children with which one is dealing is known, there is a good chance that one will be able to apply to reality one's aspirations and expectations for the children's success in "formal" school work.

As different people appreciate statistical data in different ways, an attempt will now be made to present in a variety of settings the essential facts about the distribution of intelligence throughout the population in general. It must be understood that, in what follows, an *unselected or representative sample* is being considered, and not the population of any particular school. Let us first arbitrarily divide children into bright, average, and dull; then if we agree to call intelligence quotients below 85 dull, and those above 115 bright, it will be found that approximately 64% of children will be average, and the bright and dull will consist of the top and bottom 18% respectively. In a class of thirty-two representative children this will mean that there are six dull, twenty average, and six bright.

There is, however, a good deal of difference between an intelligence quotient of 85 and one of 115. At 10 years of age, for instance, an intelligence quotient of 85 would give a mental age of  $8\frac{1}{2}$ , and an intelligence quotient of 115 would give a mental age of  $11\frac{1}{2}$ . Thus if we adopt the aforementioned standard of "average" we shall find that two children of the



same chronological age may have as much as 3 years of a difference in their mental ages and still both be in the "average" category.

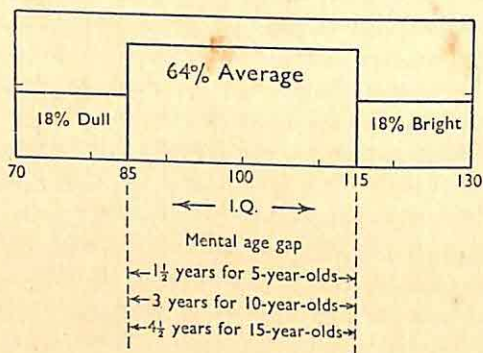


Figure II

Here again it must be emphasized that the above divisions are only arbitrary. Other limits for the group described as average might be chosen, and groups denoting degrees of dullness and brightness could be arranged. The size of these could be varied according to one's choice, though no doubt some groupings would be generally more acceptable than others.

Exception has already been taken to the above arrangement because it classes as "average" children who are very dissimilar in mental capacity. We could, therefore, narrow this range to 20 points of intelligence quotient instead of the 30 used above. "Average" would then mean intelligence quotients of from 90-110. This would mean a mental age range of 2 years for 10-year-old children, and it would also make the group smaller. We should find that only 46%<sup>1</sup> would then be classed as average, and at the same time the groups called "dull" and "bright" would be extended.

<sup>1</sup> In all these examples a standard deviation of 16.5 points of intelligence quotient is used.

We should also be labelling as "dull" all children with intelligence quotients below 90, thus grouping together as "dull" 10-year-old children<sup>1</sup> mentally retarded by one year, with others mentally retarded by three years or over. We could, therefore, split this group into the dull-to-average (I.Q. 80-90) and the dull-to-E.S.N.<sup>2</sup> (I.Q. below 80) and make new groups of average-to-bright (I.Q. 110-120) and bright-to-superior (I.Q. above 120). The size of these categories would then be 11% and 16%, and Figure II would be emended as follows (Figure III):

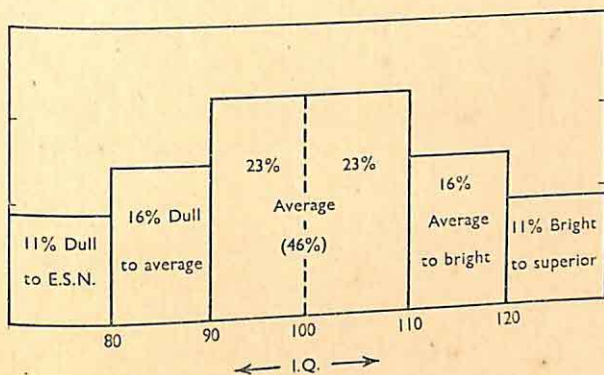


Figure III

It might be objected that it is illogical to have 20 points of intelligence quotient in the "average" group and only 10 in the others. Another diagram is therefore given (Figure IV) in which each group has an equal spread of 10 points of intelligence quotient. To these categories a set of descriptions is added. This diagram is, however, not designed to emphasize such labelling, but to call attention to a fact that may have already been noticed in Figures II and III, namely, that the size

<sup>1</sup> Reference is continually made to 10-year-old children because the corresponding mental ages work out in round figures.

<sup>2</sup> Educationally subnormal.



of the categories, or more precisely the proportion of children in each category, increases as we proceed from the extremes to the average.

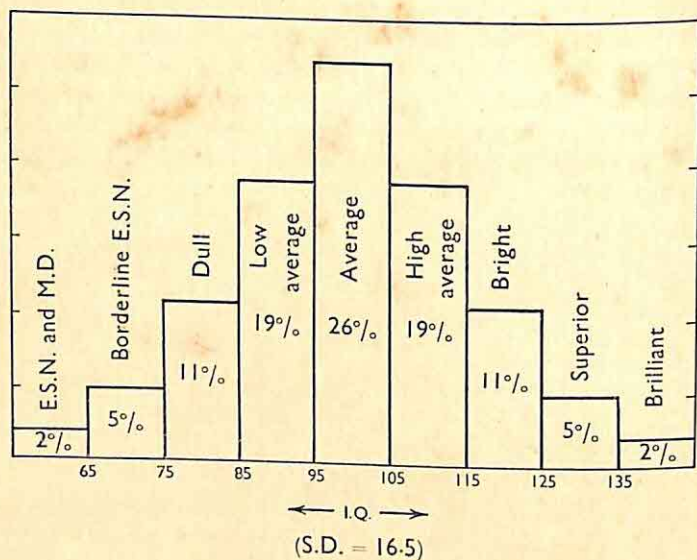


Figure IV

If still smaller intelligence quotient intervals were taken this tendency would continue, till with infinitely small intervals the outline of the histogram, as such a figure is called, would be a smooth curve shaped like a bell or hat (Figure V).

This curve is called the *normal curve of distribution* and has definite mathematical properties. It is found that many natural characteristics follow this distribution very closely, for example, man's height. Reflection will show that there are very few dwarfs or giants; most people are of more or less average height. Again, if all the children of a single age-group, for example the 8-year-olds, in a fairly large town were weighed, it would be found that when they were divided into groups

according to weight, a normal curve of distribution would follow.

There should be many occasions when Figure IV might be of use to teachers. If an intelligence test is given to any school or class, a similar histogram might be drawn when it would at once be apparent how near to "normal" the sample is, by comparing it with the shape of Figure IV. Percentages are given on this figure, these of course being the numbers in each "normal" group of a hundred; they would be halved for a group of fifty, quartered for a group of twenty-five, and so on;

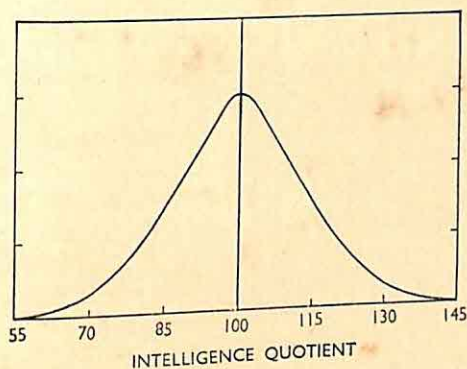


Figure V

thus in a "normal" group of twenty-five children there would be approximately one borderline educationally subnormal child (intelligence quotient 65-75; 5%), or in a "normal" class of thirty-five children there would be approximately four children with intelligence quotients between 75 and 85 (11% of 35). Here is one more example of the use of figure IV: suppose there is fair provision of Grammar School places in an area, and sixty "11+" children are an absolutely representative or "normal" sample; then (judging solely on intelligence quotient) at least four should be transferred to the Grammar School (7% of 60; intelligence quotients over 125). Some



authorities, of course, make more Grammar School provision than this for their 11+ children, and some make less. If it were known what proportion of the children were to be admitted Figure IV could be used to estimate roughly the lowest intelligence quotient likely to be needed for a child to be transferred. Thus if 10% were to receive places it can be seen that the minimum intelligence quotient would have to be about 120.<sup>1</sup> In the same way various conclusions could be reached with regard to the expected proportions of dull or borderline educationally subnormal children.

### Percentile Ranks.

The teacher who familiarizes himself with standardized tests will soon encounter the concept of *percentiles*, and if he proposes to use tests it will be profitable for him to take a little trouble to arrive at a clear understanding of this term, for it occurs frequently in Examiners' Booklets and tables of norms, and it will often be found useful in arranging a fair marking system and in conducting class or school surveys. From Figure IV it may be calculated (by adding 2% and 5%) that 7% of children have intelligence quotients below 75. An intelligence quotient of 75 would then be referred to as the *seventh percentile*. Similarly, to find out how many children have intelligence quotients below 95 one simply adds the percentage figures in the relevant columns— $2 + 5 + 11 + 19 = 37$ —and the answer is 37%. An intelligence quotient of 95 would then be the *thirty-seventh percentile*. And again, by looking at Figure IV it will be seen that there are 18% ( $11 + 5 + 2$ ) of children in the categories over intelligence quotient 115. This intelligence quotient would then be equivalent to the *eighty-second percentile* (i.e.  $100\% - 18\%$ ). From this it will be

<sup>1</sup> All the above figures are based on tests with a standard deviation of 16.5 points of intelligence quotient. In the case of a test with standard deviation of 15 points of intelligence quotient (e.g. Moray House) there would be slight differences, e.g. 5% of intelligence quotients over 125; 10% of places in Grammar Schools would give a minimum intelligence quotient of 119. See Table II.

seen that percentile marks are different from the percentage marks with which all teachers are familiar. *Percentiles refer to proportions of the population* while percentages refer to the number of marks earned out of 100.

### School Marks.

The relation between marks and percentiles can be shown in a simple form as in Table I. I have imagined a class of fifteen children and have given them percentage marks on an imaginary test.

TABLE I

CHILD	PERCENTAGE MARKS	POSITION OR RANK	PERCENTILE RANK
A	80	1	100
B	73	2	
C	68	3	
D	65	4.....	
E	61	5	75 Upper Quartile
F	57	6	
G	51	7	
H	46	8.....	
I	43	9	50 Median
J	39	10	
K	35	11	
L	32	12.....	
M	30	13	25 Lower Quartile
N	26	14	
O	20	15	
			0

Their positions or ranks in the class are shown. Then, since H is the middle child (half the children have marks higher than H and half the children are below H), his percentile is 50, though his mark is 46. Similarly, L is the middle of the lower half and D is the middle of the upper half of the group, so these have percentile positions of 25 and 75 respectively,



though the actual marks they have gained are 32 and 65. Other positions can be determined in the same way; for example, E and J are respectively 5th and 10th out of 15 and would therefore be the  $66\frac{2}{3}$  and  $33\frac{1}{3}$  percentiles.

Where form positions are determined by averaging marks in all subjects, as often occurs in terminal examinations, the marks should always first be converted to percentile equivalents. The reason for this is that teachers vary in their standards of leniency, or the examinations they set vary in difficulty, so that the top may be 90+ in subject X and only 50+ in subject Y. The top boy in subject X is given much more credit in the final average than the top boy in subject Y. There are similar wide variations in bottom marks. In a form of thirty boys, the final order between raw marks and converted marks (average of eight subjects) has been shown<sup>1</sup> to be considerable, the 2nd boy becoming 13th, the 8th boy 22nd, and the 17th moving up to 9 $\frac{1}{2}$ .

While no serious harm may come to a child who is wrongly placed in a term examination, it will readily be seen that the position is quite otherwise in a competitive examination. There is no doubt that in the past there must have been serious miscarriages of justice in respect of Common Entrance or Transfer Examinations at 11+, and it can be seen why Local Authorities should employ an educational psychologist or other competently trained officer to supervise their examination arrangements.<sup>2</sup>

### Percentiles and Standard Deviation.

It is assumed throughout this book that the reader desires to know only as much about the technical side of mental measurement as will enable him to handle tests intelligently, and to avoid falling into serious error in considering the significance

<sup>1</sup> Terry Thomas: *The Science of Marking*.

<sup>2</sup> For the purpose of combining marks for such examinations the scores should be converted to Standard Deviation Units—see Vernon or Thomas (Bibliography) or a treatise on statistics.

of test results. It has already been indicated in Chapter II that, though it is reasonable to assume that the intelligence quotient once obtained for any child will remain roughly constant, a different intelligence quotient might well be obtained—in fact might definitely be expected—if a different test were used. This is because different tests have different standard deviations. The reader is asked to take it on trust, that where

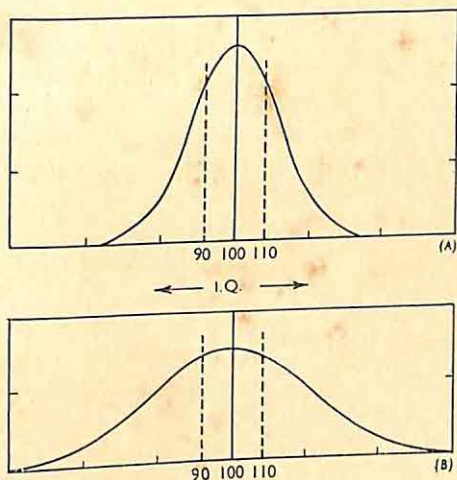


Figure VI

the standard deviation is larger there will be more children with intelligence quotients at the extremes of the scale and fewer in the centre.

This point may be more clearly illustrated by a diagrammatic representation. Figures VI(A) and VI(B) show the normal curve of distribution centred upon an intelligence quotient of 100 for two different standard deviations. Since Figure VI(A) has the smaller standard deviation,<sup>1</sup> there is a bigger area of the curve (which means a greater proportion of children) between

<sup>1</sup> Often referred to as "S.D." or " $\sigma$  score" in intelligence test literature.



intelligence quotients of 90 and 110. In Figure VI(B) this area is smaller and there are therefore fewer children within the range and, obviously, more outside it. At first sight one would take this to mean that Figure VI(B) with the larger standard deviation should produce more dull and educationally subnormal children on the one hand, and more bright and brilliant children on the other. This is obviously a mistaken approach to the problem, for a child cannot become brighter or duller merely by doing a different type of test. It is here that the superiority of measuring in percentiles becomes obvious, and Table II has been prepared with the idea of making this quite clear. The first column, headed PERCENTILES, refers to proportions or percentages of the child population, beginning with the dullest at the foot of the column. The intelligence quotients corresponding to the percentiles shown are given in columns (a), (b), and (c). This table is interpreted for column (a) as follows: .1% of the population have an intelligence quotient of 54 or less (i.e. one in every thousand); 1% have an intelligence quotient of 65 or less; 2% have an intelligence quotient of 69 or less; 3% have an intelligence quotient of 72 or less, and so on. The table can equally well be read from the upper end: thus 10% have intelligence quotients of 119 or over (since 90% have this figure or less). It is sometimes useful to approach this in another way. Suppose a child has an intelligence quotient of 113; this figure is opposite the 80th percentile, so that we know that he is better than about 80% of his fellows on a test with the same standard deviation as column (a). On the other hand, he would be surpassed on such a test by only 20% or one in five of his fellows, chosen at random from a complete sample (of the same age, of course). Referring now to column (b) and column (c) which have higher standard deviations, it will be seen that while under (a), (b), and (c) the 50th percentile in each case corresponds to an intelligence quotient of 100, this is almost the only level where such correspondence occurs. An intelligence quotient of 114, for example, in column (c) is equivalent to the 70th percentile,

in column (b) it is equivalent to the 80th percentile, while in column (a) it would be at about the 82nd percentile. This shows very clearly that the user of a test should know its standard deviation, as an intelligence quotient of 114 on a test of type (c) would be surpassed by about 30% of children, the same intelligence quotient would be surpassed by only 20% of children on a test of type (b), and by still fewer—only 18%—on type (a).

TABLE II

## APPROXIMATE INTELLIGENCE QUOTIENT

Points of Intelligence Quotient with Standard Deviation of (a) 15, (b) 16.5, (c) 27			
PERCENTILES	(a)	(b)	(c)
99.9 .....	146	151	183
99 .....	135	138	163
98 .....	131	134	155
97 .....	128	131	151
95 .....	125	127	144
90 .....	119	121	135
85 .....	116	117	128
80 .....	113	114	123
75 .....	<u>110</u>	<u>111</u>	<u>118</u>
70 .....	108	109	114
60 .....	104	104	107
50 .....	<u>100</u>	<u>100</u>	<u>100</u>
40 .....	96	96	93
30 .....	92	91	86
25 .....	<u>90</u>	<u>89</u>	<u>82</u>
20 .....	87	86	77
15 .....	84	83	72
10 .....	81	79	65
5 .....	75	73	56
3 .....	72	69	49
2 .....	69	66	44
1 .....	65	62	37
.1 .....	54	49	17

(G 906)



## DISTRIBUTION OF INTELLIGENCE QUOTIENTS

The meaning of an intelligence quotient can be misleading unless we know the standard deviation of the test used.

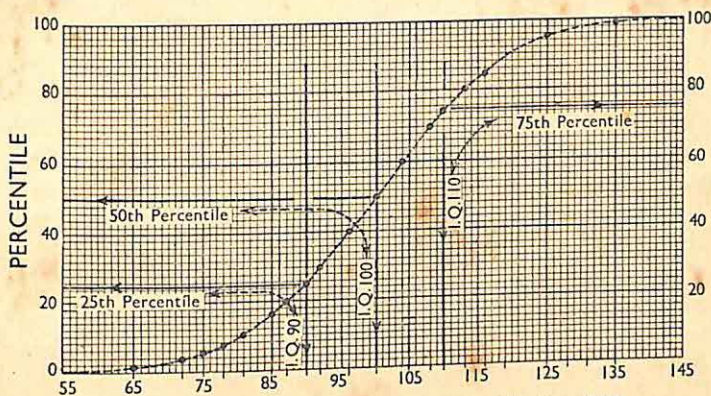
Moray House tests have a standard deviation of 15 points of intelligence quotient.

Terman-Merrill tests have a standard deviation of 16.5 points of intelligence quotient.

Some Cattell tests have a standard deviation of 27 points of intelligence quotient.

The figures shown indicate that with the Terman-Merrill test [column (b)] the 75th percentile would be 111 in a normally distributed sample, i.e. in such a sample there would be 25% above and 75% below 111. With the same test there would be 3% above and 97% below an intelligence quotient of 131 (97th percentile). We may also deduce from column (a) (Moray House tests) that only one child in a random thousand (.1%) would have an intelligence quotient above 146 or below 54.

It is almost universally agreed among psychologists that all tests would be better scored in percentiles instead of intelligence quotients, but the concepts of "I.Q." and "mental age" are so well established, and so many tests have norms published using these terms, that it will take some time to educate the test-using public in the use of percentiles. Since many tests have a standard deviation of 15 points of intelligence quotient, a percentile curve is presented in Figure VII which gives the data of Table II in graphical form [so far as column (a) only is concerned]. From this any intelligence quotient can be converted to an approximate percentile rank. Thus an intelligence quotient of 85 corresponds approximately to a percentile rank of 17. Therefore on a test with a standard deviation of 15 points of intelligence quotient, 17% of the population will have intelligence quotients of 85 or below. It will further be seen that the percentiles corresponding to intelligence quotients of 90 and 110 are respectively the 25th and 75th. This means that 50% (75%-25%) of the population have intelligence quotients between 90 and 110.



Illustrating the data of Table II Col. (a) S.D. = 15 points of I.Q.

← I.Q. →  
Figure VII

### Ideal and Actual Distributions.

It has been mentioned earlier in this chapter that the curve of normal distribution follows a mathematical law. Those non-mathematical readers whose schooldays are not too far behind them may remember "graphing" certain equations. The "normal" curve is an algebraic expression reproduced in diagrammatic form. It is thus primarily an "ideal" curve, and all the subsequent data are deduced from this ideal curve. It is legitimate to do this because in fact certain natural qualities, including intelligence, are found to be distributed throughout the population in a way which follows this curve very closely. It must never be forgotten that all the foregoing figures refer only to this ideal or total population, and school or class populations will not necessarily, in fact will rarely, follow this distribution. The intelligence quotients of all the children in an age-group in a large town or city can be expected to follow this curve very closely, and in fact do so, but those of smaller groups do not, and various tendencies, social, economic, and psychological, tend to disturb the normal distribution of intelligence in any group or geographical area.



For instance, before the Second World War, when the housing situation allowed for the easy flow of population from one area to another, it was thought that the bad living and working conditions prevalent in the rural areas were driving the more intelligent people to move towards the towns where they would find many varieties of remunerative employment. There was at any rate a certain amount of evidence to show that the mean intelligence of urban children was higher than that of children in surrounding rural areas, though not necessarily than that of children in remote country districts, presumably because the attractions of the town were not strong enough to cause an exodus of the best scholars in these areas.

Between the wars, when there was much speculative building, whole estates composed mainly of owner-occupiers sprang up. These people were frequently superior artisan or lower professional groups, of superior intelligence, whose children tended also to be of superior intelligence, so that the standard of a school population composed largely of these children would be very high, and the lower end of the normal distribution curve would be thinned out, with a corresponding bulge at the higher end. Reference was made early in Chapter I to the reverse of this problem, where an area is progressively drained of its more intelligent children when for one reason or other it becomes socially undesirable. In addition to such population movements, which are not consciously directed, there are others which cause certain school problems, and which arise because of the policies of Local Authorities. Some housing estates have been built to rehouse people from slum clearance schemes. To collect large numbers of people whose cultural level is admittedly low, and to house them, as has happened, in comparatively segregated areas, is now seen to be most inadvisable, and the people who are usually most aware of the unfortunate results of such ill-considered planning are the teachers in the schools which frequently have been built to accommodate the children in the new housing schemes. Apart from the problems arising from unleavened social standards,

these schools have to deal with those set by an almost uniformly dull school population.

Most sizeable towns have their "east-end" and "west-end" areas, and there are usually great differences between the intelligence standards of the school populations of these two types of districts. Not only are some schools full of children markedly below average intelligence, and others full of markedly above average, but others, again, situated between two dissimilar areas, and being fed by children from both, are found to have two almost distinct types of child—distinct, that is, in intelligence and social level.

At this point reference should be made to the question of the relationship between intelligence and social level. In attempting to make my argument quite clear, I have until now been considering extreme cases, but, while these do exist, and while a high social level may mean a high intelligence and vice versa, it would be unfair and misleading to *identify* absolutely social level with intelligence level. Nevertheless, there are signs that the two are associated. A few years ago I carried out an investigation with several hundred children and arranged them in groups according to the occupations of their fathers. When the *mean* intelligence quotient of each group was determined it was found that there was a progressive increase in its value as one passed from "unemployed" fathers, through "labourers", "artisans", and finally to "professional" classes.

Other investigators also have found this same tendency with larger samples, but though it undoubtedly exists, it is only the mean intelligence quotient of the group which increases, and there is a considerable overlap between classes. The child of the professional man *may* be duller than the child of the labourer, and vice versa, but on the whole this is not so.

I have also assumed that there is a correspondence between social class, occupational group, and geographical area. This assumption is only partly true, and, like all generalizations, it has many exceptions, but some readers will be able to relate the conditions of their schools to some parts of the foregoing dis-



cussion, which has been included to demonstrate some of the reasons for frequent divergencies of the distribution of intelligence away from the "normal". Consider now the mean intelligence quotient in a school. If the distribution of intelligence were normal this would, of course, be 100, but in point of fact it is easy to find schools with mean intelligence quotients of 95 or 105, mean intelligence quotients of 90 or 110 are not strikingly uncommon, while means of 85 and 115, though certainly not of frequent occurrence, are found, and 115 is exceeded here and there.

These mean intelligence quotients all refer to primary schools prior to any process of planned selection. The exceptionally high mean intelligence quotients are due, in my opinion, firstly to social factors, and secondly to a process of "natural selection". A "good" school acquires a reputation, and intelligent parents try to send their children to it, just as they tend to avoid sending them to a "poor" school if possible. If now we consider the implications of these deviations from normal distribution for individual schools it will be seen that they are considerable. Let us consider the case of a school with a mean intelligence quotient of 95. From Figure VII it will be seen that with a normal distribution this intelligence quotient would be the 40th percentile. This means that there would be 40% of intelligence quotients below 95, but in the school under consideration there would be 50%—as the *mean* I.Q. is 95. We do not know, of course, how the children spread out about the mean, but we can at once say that the school contains 10% more low average and dull children than the "normal" school, with a corresponding reduction in the proportion of brighter children. This should be a relevant factor in organizing the school. It means also that, as the average mental age of the 10-year-old group is only  $9\frac{1}{2}$  years (95% of 10 years), presumably the syllabus should be suited to children 6 months younger than the average. For schools with lower mean intelligence quotients this allowance would have to be correspondingly greater, reaching as much as a year and a half in the "poorest"

schools, while "better" schools would be able to achieve correspondingly more. The 10-year-olds in a school with a mean intelligence quotient of 115 would have a mean mental age of  $11\frac{1}{2}$  years (115% of 10 years).

So far we have been assuming something like a "normal" spread of ability about the mean, whatever it is, that is, a spread taking the general shapes shown in Figures II to V, but in a school like one of those already mentioned, drawing its pupils from two different social areas so that there are two

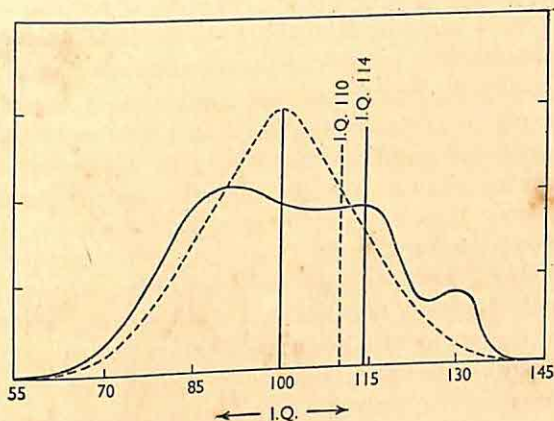


Figure VIII

"types" of child, it would be quite possible to have a distribution like that shown in Figure VIII, which would still give a mean intelligence quotient of 100. The dotted curve is the normal curve, and under this distribution there would be 25% of the children with intelligence quotients between 100 and 110, but under that of our hypothetical school the upper quartile<sup>1</sup> would move out to intelligence quotient 114. That is, the top quarter of the children would have intelligence quotients of

<sup>1</sup> The middle of the top half of a distribution is called the *upper quartile* and the middle of the bottom half is the *lower quartile*.

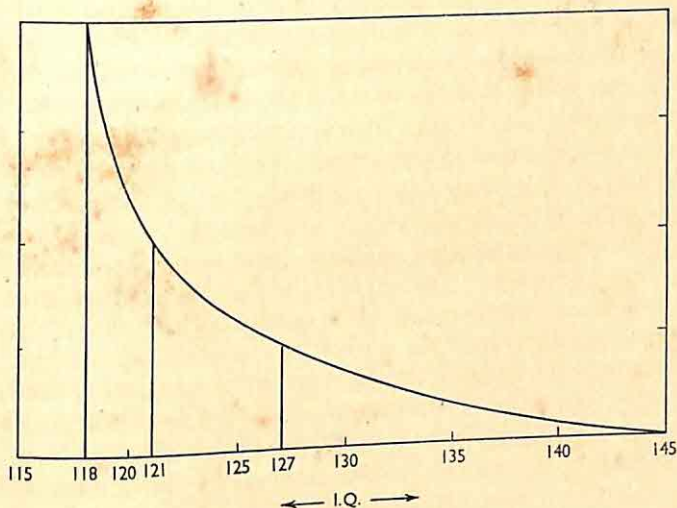


114 or over, as compared with 110 or over under a normal distribution. There would be a similar shift downwards in the lower quartile.

It will be fairly obvious that a knowledge of the distribution of intelligence in a school is very important from the point of view of school organization. In a four form entry, for instance, the median (middle) intelligence quotient and the quartiles would be rough dividing points for the four forms. In a two form entry the median intelligence quotient would be the rough dividing point for the two groups, and the quartiles would then be the middle of the two forms. The *spread* of intelligence quotient in the various forms is important. Leaving aside atypic distributions, and considering a school with more or less normal distribution, we are faced with the fact that the centre form or forms will have the narrowest intelligence quotient spread. Take for example a junior school with three streams: if we divided the streams into equal intelligence quotient ranges, say 70-90, 90-110, 110-130, we see from Figure VII or Table II that there would be twice as many children in the middle group as in each of the other two groups. With an age-group of a hundred children this would mean that there would be fifty in the middle group and only twenty-five in each of the other two, an obviously unsatisfactory grouping. There is justification, however, for making the centre group *slightly* larger than the others, and that is its homogeneity, as the following division shows. From Figure VII we see that the intelligence quotients corresponding to the  $33\frac{1}{3}$  and  $66\frac{2}{3}$  percentiles are approximately 93 and 107, so that, making the three groups equal in size, we have intelligence quotient ranges of 70-93, 93-107, and 107 to 130+, i.e. 23, 14, and 23. In terms of mental ages for 10-year-olds this means ranges of 2.3 years, 1.4 years, and 2.3 years. This assessment excludes the bottom and top 2% which would make the intelligence quotient and mental age range of the extreme groups very much larger. The teachers of the bottom and top groups therefore have a much greater range of ability to deal with, and it would

seem feasible, wherever school organization allows, to make the bottom group numerically smaller in order to narrow the mental age spread as much as possible.

The reader is reminded that the above discussion relates to the normal distribution, but, of course, distribution problems vary greatly in magnitude from one school to another. It is



Hypothetical distribution of Grammar School population  
into three equal groups according to I.Q.

Figure IX

hoped that these remarks will give point to the need for finding out the actual distribution in each school. One final example of the theoretical distribution of intelligence quotient is illustrated in Figure IX. This is the upper tail-end of the normal curve, and will fairly approximately fit conditions in many grammar schools, save that in a graph representing the population of a particular school the dividing line (intelligence quotient 118) would not be a sharp line as in Figure IX. In this diagram school population is divided into three *numerically* equal



groups, and the corresponding intelligence quotient ranges are 118-121, 121-127, 127-145. The mental age ranges corresponding to these groups would be roughly 6 months, 1 year, and  $2\frac{1}{2}$  years in the middle school.

### Distribution of Attainments.

In practice when the classification of children is being considered intelligence quotient and mental age are not the only factors which have to be taken into account. Another very important factor is attainment. This question is considered at length in Chapter VII, but it seems appropriate to mention here that, while in theory attainment distributions might match intelligence, this is rarely the case in practice. In the first place, owing to the teacher's tendency to adapt his teaching to the central level of the class, the spread of attainment is usually much less than the spread of intelligence, and this tends to make the standard deviation smaller. Variations in teaching methods, in the number of test items given in a year, and in the degree of similarity of test items in different parts of the scale, all tend to make for variations from normal distribution. The bright youngster of 7 who has a mental age of 9 is unlikely to have a spelling age of 9 unless he is an exceptionally good visualizer, as he will not have had enough writing experience to achieve this. It is thus unwise always to try to equate attainment and mental age. It is usually quite sufficient to arrange the children so that median attainment and median mental age are not too dissimilar, and then, arranging them in quartile groups for attainment and mental ages respectively, to see that there are no pronounced discrepancies. This is discussed more fully in Chapter VII.

### Rating Scales.

With the general introduction of scholars' record cards most teachers have become familiar with the idea of *rating scales*. Such scales are really an abbreviation of the percentile ranking system. The commonest is the five point rating system by

which teachers are asked to rate their pupils as "A", "B", "C", "D", or "E", in descending order, for general intelligence, ability in a school subject, or in some character quality such as perseverance. They are usually instructed to divide the group under consideration into given proportions, for example, 5% of "A"s, 20% of "B"s, 50% of "C"s, 20% of "D"s, and 5% of "E"s. This is the same sort of distribution which we have been discussing, and is clearly our old friend "normal" distribution. While the idea of rating pupils in this way is a good one if the distribution is normal, the reader who has been in any way convinced by the previous discussion, or who is aware how atypic the distribution of ability in a particular school may be, will realize that such a procedure cannot lead to an accurate representation of the facts. A school, for instance, may have no "A"s or perhaps only 10% of "B"s, while it may have 30% or more of "D"s.

The system is more suitable for a division according to character qualities, but it is very difficult to work out in practice. A teacher of forty children, asked to rank them for such a quality as "stability", has an extremely difficult task. It is not so difficult, however, to distinguish the outstanding children possessing or lacking some quality of character or temperament, and I would suggest the use of a record card on which teachers should mark, with a plus or a minus, for example, only those children possessing or lacking any particular trait to a marked degree. There is a danger in the enthusiasm for record cards that teachers may be asked to do what trained psychologists would shrink from doing. It should be pointed out that some qualities which have found their way into record card scales for assessment are probably not in any way unitary character traits. We all know the child who will concentrate for hours on something which interests him, or which he can do well, and who shows quite a different attitude to another subject, or the child who works hard for one teacher and "plays about" for another. How are we, then, to estimate such a quality as "perseverance"?



## CHAPTER VII

### CLASS AND SCHOOL SURVEYS

It is fairly easy to work up a certain amount of interest and enthusiasm for tests, but unless the data obtained can be kept in perspective and handled expeditiously the interest will soon fade. Many a teacher has compiled long lists of children's names with their marks or intelligence quotients tabulated, only to feel that he has taken a great deal of trouble to obtain a result of doubtful value. He may even feel that there is nothing constructive or effective about test results. This, of course, is quite true, and all efforts at testing are likely to be abortive if it is not kept clearly in mind that testing is principally a means of classifying, and that any particular result is valueless unless it is related to other results. Testing is an aid to making judgments. Sound judgments are the basis of effective action. The teacher, therefore, who does not proceed from test results to judgments and thence to action might as well not take the trouble to test at all.

Test results have no *intrinsic* virtue; they become useful only to those who can see an interconnection between the various results, that is, to those who are prepared to think intelligently about the data which the tests furnish. This point is stressed because, though certain suggestions are made in the following pages about handling test results, they are not the only possible ways of doing so. Moreover, since conditions in schools vary much, from the village school with one or two teachers to the urban school with from one to two hundred children in an age-group, there must be many different types of school organization, and consequently ways of arranging test data will also vary according to circumstances.

*It is especially stressed that the teacher is not being advised to take up all the following methods of handling test results. Some of them may be useful as they stand, or in part, but the lesson of the chapter is that test results must be related to one another to become useful.*

Whatever use is made of test data, however, it must always be borne in mind that a child or group is always being compared, by implication, with the "norm" or hypothetical average child or group. Moreover, so far as all attainment test results are concerned, there is no binding or imperative condition making the "norm" into the "ideal", though many teachers are prepared to accept it as such. Ignoring for the moment variations in intelligence, it may be argued that if a child or group is found to have reached a level, in arithmetic for example, which is in advance of or behind the average, the teacher may pass judgment on the position and leave it at that. Most teachers will think that the achievement of children on the whole is a reliable criterion, but attainment norms are after all partly the product of the policy of arithmetic teachers, and tests do not absolve their users from adopting their own policy with regard to curriculum content and the appropriate amounts of time and effort to be spent on the subject.

So far as intelligence is concerned, however, tests are not greatly influenced by teaching, or at least such influence should be avoided as far as possible, and therefore a teacher has not here the same right to accept or reject the norms which the test provides. Nevertheless, the reliability of all tests is such that it is not wise to expect of them a hair-splitting degree of accuracy. Making the necessary allowance for this, we can then say with regard to results in general that the intelligence result is the basic item to which attainment results have to be related, but not necessarily equated, though most teachers feel it is reasonable to aim at such a level of attainment as does in fact match the intelligence level.



### Individual Records.

It has been urged that individual graphical records should be kept of each child's scholastic progress, in conjunction with his developing mental age. This would be very interesting but would perhaps take up more time than many teachers feel they can afford. The following illustration of this type of

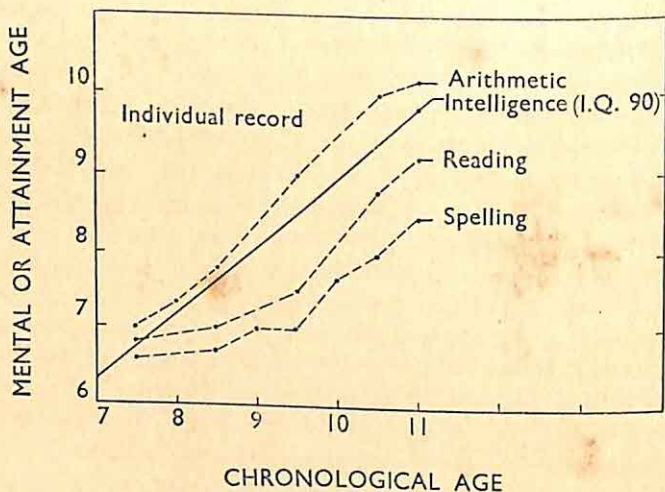


Figure X

record is given (Figure X) as it shows in simple form a method, to be discussed later, of dealing with class or school results in a similar way.

It is seen from an intelligence test that a child has an intelligence quotient of 90. At 7 and 11 years respectively he will have mental ages of 6.3 and 9.9 years (90% of 7 and 11 respectively). These two points are marked on the graph and joined with a straight line, it being assumed that the intelligence quotient is constant, and that the ratio of mental age to chronological age is therefore constant. Suppose now that he is given standardized tests at intervals, and that his attainment ages

for each test are marked. At  $7\frac{1}{2}$  years his attainment ages are approximately equal to his mental age. Arithmetic is his best subject and spelling his worst: he likes arithmetic, and while he is always below average in this subject he makes such progress that his arithmetic age is always a little in advance of his mental age. Let us suppose that he has been transferred from a small infant department, where he has had a good deal of individual attention, to a junior department, where the classes are large, and where there are many poor readers. He cannot have the kind of individual help to which he has been accustomed, and has an unsuitable reading-book. He becomes discouraged, so that he loses his zest for learning, and at the end of the year it is found that he has made very little progress. He does a little better than this in the next year; in the third year he has a teacher he likes and with more help improves rapidly. In spelling the position is similar to that in reading, though this is his worst subject all through the school. Tests are marked at six-monthly intervals. In one period he makes no progress at all. This may be a true representation of his spelling progress, but, on the other hand, it may have been a bad day for him when he was tested. The graph shows, however, that he is approximately 18 months retarded in spelling from  $9\frac{1}{2}$  to 11 years, an amount which most teachers would regard as serious. Suggestions for tackling this problem are made elsewhere in this book.

### Class Records.

A record similar to that shown in Figure X can be kept to show the average progress of a class. Here the intelligence quotient, used as a reference line for the attainments of the class, would be taken from the average of all the intelligence quotients of the class. The attainment ages marked from time to time would be the average of the attainment ages, in the respective subjects, for the class, and these would be marked off at points corresponding to the average chronological age of the class.



## Median and Quartiles.

The calculation of averages takes much time, but averages were mentioned in the foregoing paragraph because they are well understood. A much less laborious way of finding the central tendency in a group is to use the *median* or middle score. This usually differs little from the average, and in some ways is superior to it. To find the median score one has only to arrange the test-papers in ascending or descending order and count to the half-way mark. The score at that point is the median score. In the case of an even number of children there will be two middle scores, and a point midway between these has to be taken. Thus in the case of twenty children ranked in order, the tenth and the eleventh are the middle ones, as there are nine above and nine below them. If the mental ages of the tenth and eleventh children were respectively 9 years and 9 years 2 months, the median mental age would be 9 years 1 month. From such a graph, which would be of the same type as Figure X, it could be seen at a glance how the various attainments compare with the mental age of the class as a whole.

There is a defect in this system, however, because in taking the median (or the average) to represent the class we have no idea of how the children are spread out, and the same median (or average) might be found if the difference between the top and bottom children were one year or five years of mental or attainment age.

Some idea of the spread of ability or attainment can be gained by plotting also the *quartile* scores. These are the scores of the children who come in the middle of the top and bottom halves and are known as the upper and lower quartile scores respectively. Thus in a class of forty children ranked in order of their scores the twentieth and twenty-first would be in the middle, and these would give the median score. There would then be nineteen children above and nineteen below the median. The score of the tenth child (middle of the top nineteen) would be the upper quartile and the score of the thirty-first child

(middle of the bottom nineteen) would be the lower quartile.

Table III shows how twenty-nine children might be found to be arranged and how the median and quartiles would be determined.

TABLE III

CHILD	RANK	MENTAL (OR ATTAINMENT) AGE	
A	1	10 years 2 months	
B	2	..... etc. ....	
C	3	.....	
D	4	.....	
E	5	.....	
F	6	.....	
G	7	8 years 8 months	} 8 years 6 months
H	8	8 years 4 months	
I	9	.....	
J	10	.....	
K	11	.....	
L	12	.....	
M	13	.....	
N	14	.....	
O	15	6 years 6 months	
P	16	.....	
Q	17	.....	
R	18	.....	
S	19	.....	
T	20	.....	
U	21	.....	
V	22	6 years	} 5 years 11 months
W	23	5 years 10 months	
X	24	.....	
Y	25	.....	
Z	26	.....	
AA	27	.....	
BA	28	5 years 1 month	
CA	29		

It should be clearly understood that in practice it would not be necessary to make out such a table. The median and quartile

(G 906)



scores could be found by merely counting down the marks on the test papers when these have been arranged in order.

Figure XI shows the concept of median and quartile scores applied to reading and mental ages for a class as it passes through the junior department. The quartile and median lines are drawn by joining the mental ages corresponding to two convenient chronological ages. Thus in Figure XI

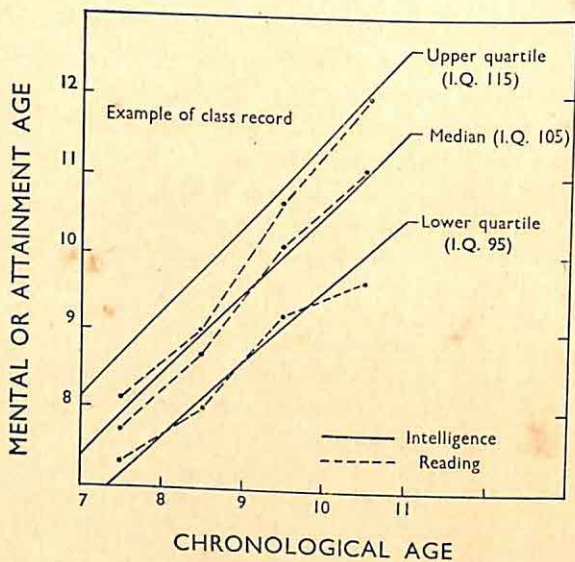


Figure XI

the lower quartile for the class is an intelligence quotient of 95. From this we calculate mental ages of 6.65 and 10.45 years at 7 and 11 chronological years respectively (i.e. 95% of 7 and 11 respectively). The lower quartile line then joins 6.65 years mental age at 7 years chronological age to 10.45 years mental age at 11 years chronological age. The same method is used to draw the median and upper quartile lines. It is assumed that the children's intelligence quotients remain

constant, and that the mental ages therefore increase at a steady rate. In contrast to this the attainment scores fluctuate. Half the children are included between the upper and lower quartiles. It will be seen that at  $7\frac{1}{2}$  years the reading ages of the middle 50% are much more compressed than the corresponding mental ages. This, presumably, is owing to the teaching being directed to a central level in the class. Suppose the median intelligence quotient is 105. The median reading age is a little better than average, but not quite so good as the median mental age. The reading lower quartile is slightly better than the corresponding mental age quartile, while the reading upper quartile is considerably inferior to the corresponding mental age quartile.

Suppose during the next year the children have only mediocre teaching. The position is then as shown at  $8\frac{1}{2}$  years. Subsequently the children have some good teaching, and the teacher realizes that the brighter children can progress faster than the rest. At  $9\frac{1}{2}$  years all the children have made excellent progress, especially the brighter ones, and the spread of reading ages is much more in line with the mental age spread. During the next year this progress is maintained for the average and bright children, but the duller children have fallen off a little. This tendency would have to be watched.

It should be clear from Figure XI that if the spelling and arithmetic data were added the figure would become very complicated. It is therefore better to have separate diagrams for each subject. This involves drawing the mental age median and quartiles as reference lines for each subject unless graph tracing paper is used, and in this case they need be done only once, as the diagrams for each subject can be superimposed on this. This method has the additional advantage of facilitating the comparison of the various subjects with each other.

This system of keeping a class record will be found satisfactory only where the class has a relatively stable population. Where there are frequent changes, because, for example, of the transferring of children from an A to a B class or vice versa, it



cannot be assumed that the median and quartiles for mental age would be straight lines. The changes would have to be taken into account and a new median mental age found, with the corresponding quartiles each year. This should not be difficult if all the children's intelligence quotients are known, for they could be ranked in order, and median and quartile intelligence quotients found by counting through. It would then be necessary to find the median chronological age of the class and calculate the median and quartile mental ages corresponding to this.

For example, suppose a two stream secondary school has a B class with quartile and median intelligence quotients of 110, 95, and 78. Some attainment tests are given at a time when the median chronological age of the class is 11 years 7 months. The attainment ages will then have to be related to the mental ages *at that time*, so that we must find 110%, 95%, and 78% of 11 years 7 months to reach the quartile and median mental ages of the class. These will be 12 years 9 months, 11 years, and 9 years 4 months. Suppose then at the end of the year there is some shuffling of the children, so that the brightest go into the A class and some of the dullest of the A class move down into the B class. The quartile distribution of the B class now becomes more homogeneous, and we find the three critical intelligence quotients have become 102, 92, and 77. If an attainment survey is then to be carried out we must first find the median age of the class and then the distribution of mental ages. Let us suppose that the median age is 12 years 10 months: 102%, 92%, and 77% of this age are respectively 13 years 1 month, 11 years 9 months, and 9 years 11 months; these are then the quartile and median mental ages which will act as reference points for the attainment age distribution. By comparing the quartile distributions for the two years, it will be seen that they have been increased by only 4 months, 9 months, and 7 months respectively, so that the work of the class could be expected to improve only by these amounts approximately.

*The position of individual children* should be considered in relation to the class median and quartiles, whether such a graphical record as Figure XI is made or not. A child ranking above the upper quartile for mental age would normally be above that quartile in attainment. If such a child were found below the bottom quartile in one of his subjects it would clearly present a case for investigation. Again, if a child were found to have attainment scores above the median, but had a mental age in the bottom quarter, one might feel justified in looking into his intelligence test score; it may have been unreliable, or he may be one of those children of stable temperament who do unexpectedly well at school.

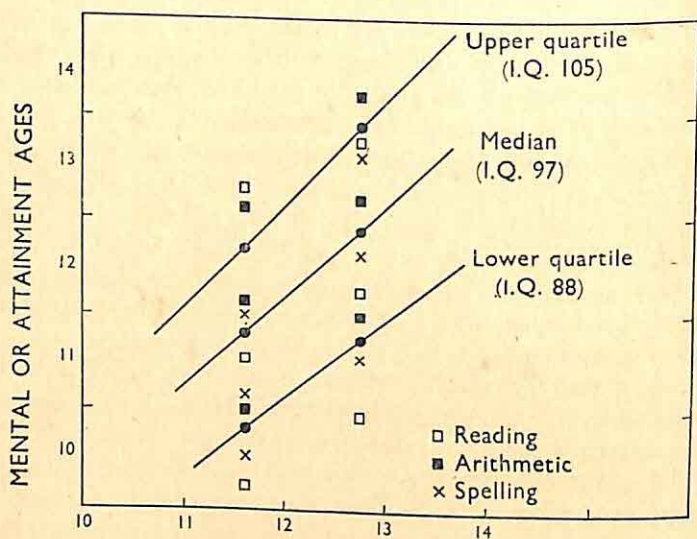
### Age-Group Records.

This example of classes being regrouped after a period raises the whole question of the relationship between the classes in an age-group. Much interesting information can be obtained by comparing the records of the A, B, and C classes as already described, but greater perspective would be gained if a similar record were made for the complete age-group, i.e. all the A, B, and C children combined. This would be of particular interest to the head teacher, for he would be able to see: (a) whether the group was normal and normally distributed for intelligence (if it were not he could consider modifications in the syllabus); (b) whether the attainment and mental age of the group, both as a whole and as at the upper and lower quartiles, were keeping in step (this would help him to judge whether there was any weakness in coping with a certain section of the distribution); (c) whether a proper balance was being kept between the various subjects; (d) whether the progress from year to year was satisfactory.

Figure XII shows an imaginary age-group in a Secondary Modern School. The children in it are tested in their first and second years at the school, and the various medians and quartiles are obtained as shown. It will be seen that the arithmetic



is uniformly good, and this position is maintained after the children have moved up to the second year. The spelling is not so good in the first year but improves by the time the children are tested again, especially in the case of the average and bright children. The reading in the first year is both good and bad; the best children are clearly excellent but the poorest



Intelligence and attainment spread for an age-group

Figure XII

ones are not as good as their mental age might suggest. In the second year the reading has deteriorated at all levels, and though progress has been made there has been a falling away both in relation to mental age and to other subjects. It might be mentioned that where diagrams tend to become a little complicated, the use of a larger scale, and coloured inks or pencils for the various subjects, restores the clarity of what is being portrayed.

### Age-Group Spread.

Once a teacher has become accustomed to noting where a child stands in relation to the median and quartiles he will have found the key to much valuable information. Beyond the upper and lower quartiles it is usual to find a gradual tapering off, as is shown in Figure IV (the normal distribution). Sometimes, however, a bunching of children's scores may be detected and a clearer picture of this situation may be desirable. Where this is so a histogram will be found useful. Figure XIII is such a diagram drawn to show a fairly

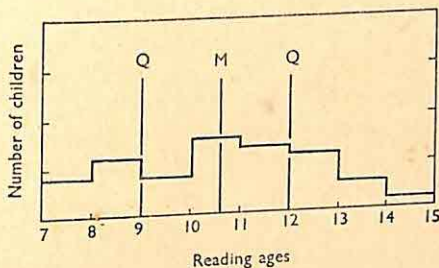


Figure XIII

normal distribution of reading scores except for a bunching of poor readers. It is not uncommon to find an age-group with an attainment spread as large as this in a Secondary Modern School where the children have come from a variety of contributory junior schools. Figure XIII demonstrates very clearly the need for "streaming" or classifying the children to make teaching in large classes at all practical.

This process of "streaming" must also be considered from the point of view of intelligence quotient or mental age. Figure XIV shows an example of poor grading. If the three groups were put together the sample would be a normally distributed one, and if it were divided into three numerically equal groups according to intelligence quotient we should have C—below 65 to 92; B—93 to 108; A—109 to 135+. It can be seen that



the grouping in Figure XIV is very far from this ideal division. Even to one without a knowledge of the theoretical dividing lines it is fairly clear that there are children in the A group who should be in the B or C group. Moreover, there is not a great deal of difference between the B and C groups save that group C certainly contains the duller children. One-third of group C has intelligence quotients of 95-105, and these children, being "average" children, should certainly be in the B group, unless some very special circumstances make it advisable to keep one

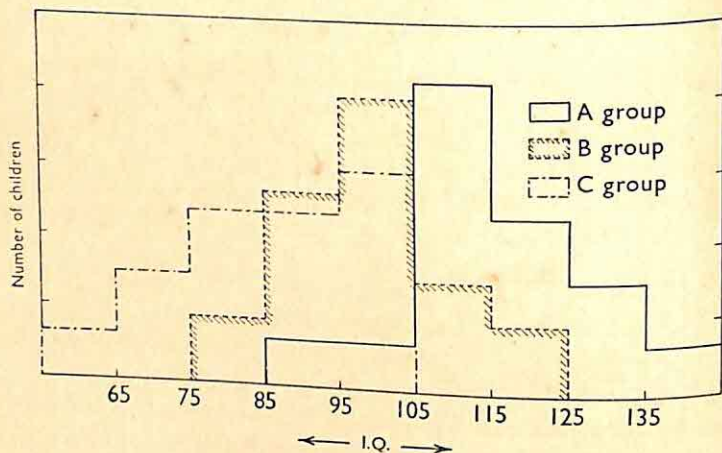


Figure XIV

or more of them in the C group. It must, of course, be made quite clear that in practice it would not be sensible to draw clear-cut lines at intelligence quotients of 92 and 108 to divide such a group. Such lines might be considered as guides, but other factors would have to be taken into account. For teachers who need guidance in making a reasonably scientific success of directing children into appropriate streams some procedure such as the following might be adopted. It is drawn up for use with three streams but could be modified for a different number of groups.

Suppose an intelligence test and three attainment tests have been given. Then proceed as follows:

1. Arrange the test papers in order of ascending mental (or attainment) age. In each case divide them into top, middle, and bottom thirds.
2. Note all the children in the top third for mental age who also have all three attainment ages in the top third of the respective subjects. These should comprise the core of the upper group.
3. Invert this process for the lowest third. Assuming that the three groups are to be equal in size, it will then be known how many short of the requisite number the highest and lowest groups are.
4. Continue to fill the top group, up to the required number, with children in the top third for mental age, and also in the top third in *two* attainment tests.
5. Invert this process for the bottom group.
6. If the numbers in the top group are not made up by this procedure, consider as suitable for this group any child with attainment ages in the top groups for each subject, who has a mental age nearly equivalent to those at the bottom of the top third of the mental ages.
7. Similarly with the bottom group, those whose attainments clearly place them in this group, and whose mental ages are near the bottom of the middle group, would be suitable for moving into this category.
8. There would then be few, if any, children left about whom individual decisions would have to be made. Clearly some at the border-lines could be suitably placed either above or below.

The above procedure is, of course, fairly mechanical. There may be factors other than the test results to be taken into account, and it is not suggested for one moment that a mechanical process of this kind should be followed with such rigidity that it would produce an unwise decision in the case of any particular child.



### School Surveys.

So far we have considered only classes and age-groups and the help to be gained from tests in the organization and recording of their progress. A separate record would have to be kept for each class or age-group throughout the school. Such records would have data added to them at convenient intervals, perhaps once or twice a year, and would become complete only when the groups of children with which they were concerned had passed right through the school. Though it is clear that such a system would cover the whole school it might not give a composite picture of the whole school at any one time, but only a number of diagrams of the various age-groups or classes, each at various stages of progress through the school.

It is, however, also possible to take the data from these age-group records and reassemble it on a single diagram to give a static picture or a kind of cross-section through the school *at a particular time*. This point must be emphasized because, though the appearance of the diagrams reproduced below is very similar to that of earlier diagrams (Figures X to XII), they are in fact quite different in their purpose. Moreover, to obtain a static picture of the school at a given time it is not essential to produce age-group records first, and then abstract the data from them. Head teachers may wish to survey their schools without wishing to make cumulative class records, and this, of course, can be done quite independently.

If this were the case all the children in the school would be tested and the results recorded on a graph such as that shown in Figure XV. To make the position as simple as possible let us suppose it is required to survey the state of the mechanical arithmetic throughout a school at a particular time. The median and quartile scores are marked for the 7, 8, 9, and 10-year-olds. These scores are marked off at the median chronological age of the group at the time of the test. The test norm is also marked as a reference line. (The one given is in fact

Cattell's "Midland" norm; it is the line representing the mean score of the standardization sample.) Figure XV differs from similar diagrams given earlier in that marks are substituted for arithmetic ages. Each child could, of course, have had his score converted to an arithmetic age, and the median and quartiles could then have been found, but if the aim is only

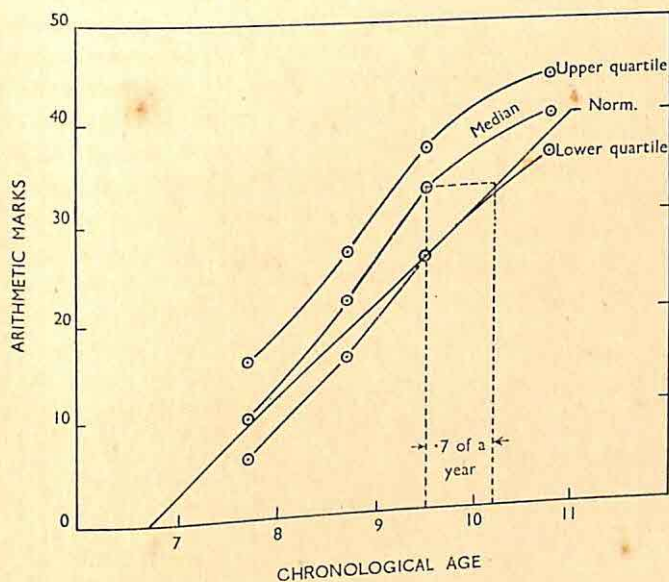


Figure XV

to survey the school as a whole, this would be unnecessary. The following comment could be made on Figure XV. The mechanical arithmetic in the school is, on the whole, superior. The 9-year-olds are mostly in advance of the norm—0.7 of a year to be precise—as the graph shows. Whether this relative superiority is due to an enthusiastic teacher, or whether the children in that year are relatively brighter, would have to be decided by an intelligence test. This test would also show



whether the general high level on the arithmetic test were due to a relatively bright sample of children in the school, or whether it were due to an emphasis on mechanical arithmetic in the school.

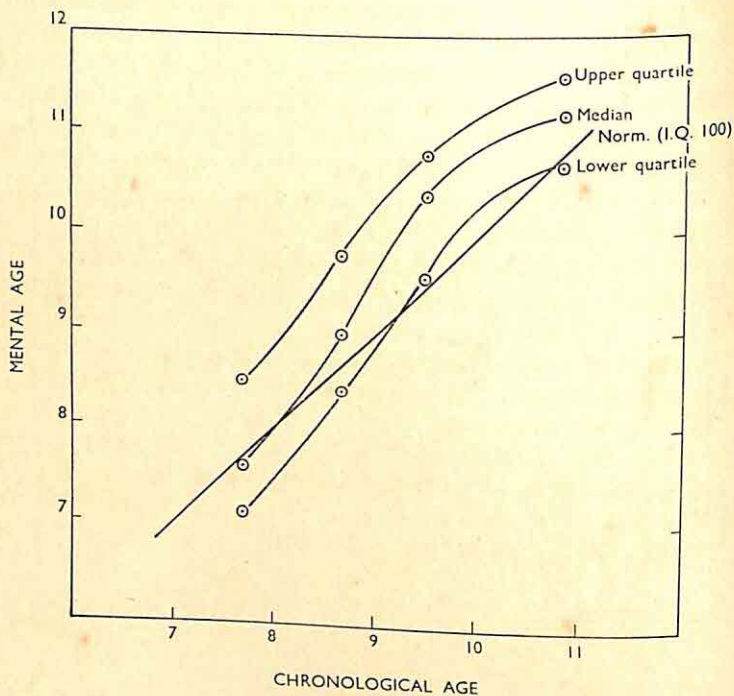


Figure XVI

Similar diagrams might be prepared for the results of other attainment tests and it would then be clear how the various subjects compare with one another. It is fairly common to find that one subject has unwittingly been neglected or, equally unwittingly, has been stressed, when such comparisons are made. A diagram like Figure XV, however, has only one reference line, the test norm; the school median can

be compared with this but it gives no indication as to whether or not the spread between the quartiles is appropriate. A rough idea of what this should be can be obtained from an intelligence test diagram. If mental age were plotted against chronological age for the same children as are concerned in Figure XV, the mental ages being found at the same time as the arithmetic ages (i.e. giving the same median chronological ages as in Figure XV), a graph such as is shown in Figure XVI might be obtained. This follows the general shape of the arithmetic results, and shows that the 9-year-olds are in fact brighter than the other groups, which would explain the higher arithmetic scores.

Figures XV and XVI are then static pictures or cross-sections of a school at a particular time. From them we know that bright 9-year-olds will in the next year be bright 10-year-olds. If, however, it were desired to follow the progress of such a group year by year, the method of Figure XI or XII would be appropriate.

### Intelligence Quotient Tables.<sup>1</sup>

Where many intelligence quotients have to be worked out from mental and chronological ages, time can be saved by using intelligence quotient tables. These tables also give mental ages corresponding to chronological ages when the intelligence quotient is known. This is very useful where a new mental age has to be found, after a passage of time, from an intelligence quotient previously determined, as when it is desired to compare attainment ages with mental age.

### Attainment Norms.

These are often given in the form:

Age	—	8	9	10	11	12
Score	—	7	14	21	27	32

<sup>1</sup> See Inglis (Bibliography).



This gives no intermediate points but these can be quickly obtained if "score" is plotted against "age" on graph paper. The Norm on Figure XV is a line obtained in this way (though such lines are not always straight lines). From Figure XV it can be seen that a score of 26 gives an attainment age of 9.5 years and a score of 40 gives an attainment age of 11 years. On a graph drawn for the table given above a score of 30 would give an attainment age of 11.6 years. If the attainment ages were preferred in years and months instead of decimals, twelve squares would have to be allotted to each year instead of ten.

I hope the reader will have discovered in this chapter something of value which he can put into practice. I should like to reiterate that all the survey methods described are not necessarily applicable to every particular situation. I know that the very sight of a graph will discourage some people but I hope that where this is so it will not cause anyone to reject the fundamental position underlying what has been said. A careful study of the information assembled on a graph will reveal that the latter is not entirely necessary. The same information could easily be recorded in figures. The essential thing is to link up test data whether by simply relating mental age to attainment age, or by more complex means such as relating the distribution of ability in a school to that of normal distribution. I am not concerned to press for the adoption of any cherished devices of my own, but I *am* concerned that test results should be given sufficient thought and consideration to make them of value. They should always ultimately benefit both teachers and children, and if they fail in either of these aims they fail absolutely.

## CHAPTER VIII

### TEST MATERIAL

Note: The volumes in which the test material is to be found are given, but where it is also obtainable separately from the publisher the test is marked with an asterisk. Where the reference is to a book the publisher can be found under "Bibliography" at the end of this chapter. Where the test material is supplied separately by the publisher it is absolutely necessary to consult the volume from which it is taken, unless the published test has a booklet of instructions and norms with it.

#### I. Attainment Tests

Note: Tests with a narrow age range, such as tests for the 11+ group, have been omitted.

#### READING.

A. WORD RECOGNITION. (All the following are oral individual tests.)

Burt: Graded Vocabulary Test. Ages 4-15.

*Mental and Scholastic Tests.*

Schonell, F. J.: Graded Reading Vocabulary Test.\* Ages 5-15.

*The Psychology and Teaching of Reading.*

Vernon: Graded Reading Test. Ages 5-adult.

Revision of Burt's Test (for Scottish children).

This book contains material and instructions for both tests in a handy form.

*The Standardization of a Graded Word Reading Test.*



## B. CONTINUOUS PROSE READING.

Burt: Test 5. Ages 7-14.

*Mental and Scholastic Tests.*

Schonell, F. J.: "My Dog" Test.\* Ages 6-9½.

*Backwardness in the Basic Subjects.*

Both the above are individual tests and measure speed, accuracy, and comprehension, giving separate measures of each. Burt's test will not measure speed or accuracy satisfactorily above 10 years.

## C. COMPREHENSION. (All the following are group tests. Times are given when in excess of 20 minutes.)

Ballard: Silent Reading (B).\* Ages 9-14.

*The New Examiner.*

Brighton Reading Tests. Ages 9-14.

6 separate tests, 40 minutes.

(University of London Press.)

Burt: Northumberland Test II. Ages 7-14.

A composite English Test, 50 minutes.

(University of London Press.)

Cattell: Midland Test.\* Ages 6½-14.

*A Guide to Mental Testing.*

Fleming: Kelvin Measurement of Reading Ability. Ages 8-12.

(Gibson.)

Schonell, F. J.: Silent Reading Test A.\* Ages 7-11.

Silent Reading Test B.\* Ages 9-13.

*Backwardness in the Basic Subjects.*

Watts: Sentences for a Reading Scale. Ages 6½-11.

Reading Tests for Seniors (4 tests). Ages 11½-14½.

*The Language and Mental Development of Children.*

## OTHER ENGLISH TESTS.

Ballard: English Construction Test.\* Ages 10-14.  
*The New Examiner.*

Burt: Composition—Median samples. Ages 7-14.  
*Mental and Scholastic Tests.*

Cattell: Midland Attainment Test—Grammar, Style.\*  
 Ages 6½-14.

Knowledge of Literature.\* Ages 6½-14.  
 (University of London Press.)

"Dingwall" Tests in English Usage. Booklets 1 and 2:  
 ages 9-10. Booklets 3 and 4: ages 10-12.  
 (University of London Press.)

Schonell, F. E.: Diagnostic English Tests. Ages 9½-16.  
 Punctuation, English usage, vocabulary  
 and sentence structure.  
 (Oliver and Boyd.)

Schonell, F. J.: Composition—Median samples. Ages  
 7½-13½.

For narrative, reproductive, imaginative,  
 and explanatory compositions.

Tests of Analysis and Synthesis of  
 Common Phonic Units.

Test of Directional Attack on Words.

Visual Word Discrimination Test.

*Backwardness in the Basic Subjects.*

Watts: I. Vocabulary Tests:

One Hundred Common Names

" " " Class Names

" " " Verbs

" " " Adjectives List A

" " " Adjectives List B

" " " Ages 10-15.

(Obtainable separately from U. L. P.)

(G 906)



Vocabulary Test for Young Children. Ages  $4\frac{1}{2}$ -8.  
Words with More than One Meaning. Ages  $11\frac{1}{2}$ - $14\frac{1}{2}$ .

Ideational Addition. Ages  $11\frac{1}{2}$ - $14\frac{1}{2}$ .

II. Sentence Construction Tests:

Sentence Completion Test. Ages 10-13.

English Language Scale. Ages 4-10.

Sentence Patterns. Ages 8-12.

Connecting Words and Phrases. Ages  $8\frac{1}{2}$ - $10\frac{1}{2}$ .

Norms derived from bright children.

Reported Speech Test. Ages 9-11.

III. Sentence Arrangement Tests. Ages  $9\frac{1}{2}$ - $13\frac{1}{2}$ .

IV. Time-Relation Tests. Ages  $10\frac{1}{2}$ - $14\frac{1}{2}$ .

V. Personal Qualities Tests. Ages  $10\frac{1}{2}$ - $13\frac{1}{2}$ .

Standardization doubtful for all except the five vocabulary tests under I, though norms are given.  
*The Language and Mental Development of Children.*

## SPELLING.

Ayre: Spelling Scale.

Consists of 1000 common words arranged in groups of ascending order of difficulty. Scored in "grades" and "standards". A useful list but a cumbersome test.

Ballard's *Group Tests of Intelligence*.

Burt: Graded Vocabulary Test. Ages 5-14.

*Mental and Scholastic Tests.*

Cattell: Midland Test 3. Ages 6-14.

A very brief test.

*A Guide to Mental Testing.*

Fleming: Kelvin Measurement of Spelling Ability.  
Ages 8-12.  
(Gibson.)

Schonell, F. J.: Test 1A. Irregular Words.  
Test 1B. Regular Words.  
High diagnostic value.  
Graded Dictation Tests. Ages 7-13.  
These show up errors under more  
natural conditions of spelling.  
*Essentials in Teaching and Testing Spelling.*

## ARITHMETIC.

### A. MENTAL:

Ballard: One-minute Oral Addition Test.\* Ages 5-12.  
One-minute Oral Subtraction Test.\* Ages 5½-13½.  
Serviceable only up to about 8.  
*The New Examiner.*

Burt: Graded Oral Test. Ages 4½-14.  
*Mental and Scholastic Tests.*

Schonell, F. J.: Test 12. *Diagnostic Arithmetic Tests.*  
(Oliver and Boyd.)

### B. MECHANICAL SKILL.

Ballard: Mechanical Arithmetic.\* Ages 9-14.  
A mixed test involving knowledge of methods  
as well as basic number combinations and the  
"four rules".  
*The New Examiner.*

Three Minute Tests.\*  
Four rules, each three minutes.  
*Mental Tests.*



Burt: Tests 11-14. Four Fundamental Rules. Ages 6-14.  
Four five-minute tests.

Norms for Scottish children from Dr. W. B. Inglis, Moray House Training College, Edinburgh.

Graded Written Test 9. Ages 7-14.

Involves knowledge of methods as well as basic number combinations.

*Mental and Scholastic Tests.*

Cattell: Midland Test.\* Ages 7-14.

Four fundamental rules. Takes twelve minutes.

Useful for a quick survey.

*A Guide to Mental Training.*

Fleming: Beacon Arithmetic Diagnostic Tests.

For use in connection with *The Beacon Arithmetic Teachers' Manual*.

(Ginn and Co. Ltd.)

Hill: Southend Attainment Test.\* Ages 6-14.

*The Education of Backward Children.*

Schonell, F. J.: Diagnostic Arithmetic Tests. Ages  $6\frac{1}{2}$ - $14\frac{1}{2}$ .  
*The Diagnosis of Individual Difficulties in Arithmetic.*

Essential Mechanical Test. Ages 7-12+.

Form A and B.

Two comparable forms of equal value—  
alternate forms can be used to avoid  
coaching effect if used repeatedly.

(Oliver and Boyd.)

#### C. PROBLEM, REASONING, OR METHODS TESTS.

Ballard: Arithmetical Reasoning.\* Ages 9-14.

*The New Examiner.*

Burt: Graded Written Test. Problems. Ages 7-14.

*Mental and Scholastic Tests.*

Cattell: Midland Knowledge of Methods Test.\* Ages 7-14.

Also useful as an oral test.

*A Guide to Mental Testing.*

#### D. COMPOSITE TESTS.

Burt: Northumberland Standardized Test I. Ages 7-14.  
Seven sub-tests, each with separate norms. Covers mental, mechanical, and methods.  
(University of London Press.)

Fleming: Kelvin Measurement of Arithmetic Ability. Ages 7-12.  
(Gibson.)

Schonell, F. J.: Essential Problem Test. Ages 8-14+. Form A and Form B.

Two equivalent tests for alternate use if there is any danger of coaching effect with repeated use.

(Oliver and Boyd.)

#### HANDWRITING.

##### A. CURSIVE.

Burt: Speed of Writing (Test 15). Ages 6-14.  
Quality of Writing (Test 16). Ages 5-14.  
Judged by median samples.

*Mental and Scholastic Tests.*

Cattell: Midland Test. Ages 13-15.

A five-point scale; norms for boys and girls.

*A Guide to Mental Testing.*

##### B. SCRIPT.

Ballard: Script Writing Scale.

An arbitrary scale; samples for marks 1 to 10.

*The New Examiner.*



## TESTS IN OTHER SUBJECTS.

Ballard: Algebra Test. Ages 11-14.

Geography Test. Ages 11-14.

History Test.

(Standardized type of test but no norms given.)

*The New Examiner.*

Burt: Northumberland Standardized Test II. Ages 7-14.

This is styled an English test, but it also contains history and geography tests with separate norms for each.

(University of London Press.)

Crossley: New-Type Tests in Chemistry.

Not standardized.

(University of London Press.)

Walton: Geometry Attainment Test.

Norms account for boys and girls, age, and period of study.

(University of London Press.)

Watts: English History Test. Ages 11-15.

Pen portraits of outstanding personalities in English history.

*The Language and Mental Development of Children.*

## II. Intelligence Tests

The ages given in this section refer to *mental* ages. See Chapters III and IV before choosing a test.

### (a) INDIVIDUAL TESTS (including performance tests).

Alexander: Performance Scale. Ages 8-18.

Also useful for deaf children.

(Nelson.)

Burt: Drawing a Man Test. Ages 3-14.

Actually a drawing test, but useful as a supplementary test for differentiating educationally subnormal children.

Graded Reasoning Tests. Ages 7-14.

*Mental and Scholastic Tests.*

Cattell: Scale O. Dartington Scale. Ages 4-8.

Non-verbal material and some oral questions.  
(Harrap.)

Drever and Collins: Performance Tests of Intelligence.

Scale A—ages  $5\frac{1}{2}$ – $15\frac{1}{2}$ .

Scale B—ages 4-7.

Instructions in *Performance Tests of Intelligence*.

Also useful for deaf children.

(Baird, Lothian Street, Edinburgh.)

Goodenough: Drawing a Man Test. Ages 3-13.

*Measurement of Intelligence by Drawings.*

Porteus: Maze Test. Ages 3-14.

Burt's *Mental and Scholastic Tests*.

Raven: Progressive Matrices. Ages 6 to adult.

Non-verbal. Scored in percentiles. Gives only an approximate grading.

(H. K. Lewis, Gower Street, London.)

Seguin-Goddard: Formboard. Ages  $3\frac{1}{2}$ –20.

Of doubtful value beyond a mental age of 6 or 7.

Instructions in Cattell's *A Guide to Mental Testing*.

(Experimental Instruments Co., Sudbury, Suffolk.)

Terman and Merrill: New Revised Stanford-Binet Test of Intelligence. Ages 2-23. Forms L and M.

(Harrap.)



Valentine: Intelligence Tests for Children. Ages 2-15.  
 The small amount of apparatus required for  
 this test can be made at home or in school.  
 (Methuen.)

(b) GROUP TESTS. VERBAL.

Alexander: Thanet School Aptitude Tests. Ages  $6\frac{1}{4}$ -16.  
 (University of London Press.)

Ballard: Columbian Mental Test.\* Ages 10-14.

Chelsea Mental Test.\* Ages 11-14.

*Group Tests of Intelligence.*

Group Test for Juniors. Ages  $6\frac{1}{2}$ -13.

Oral test; answers written.

*The New Examiner.*

According to norms given in *Notes on the Record Card* by the Kent Education Committee, Ballard's norms give mental ages that are too high for the lower ages. The norms are also extended in the *Notes* to 6 and 7 years, for which the scores are respectively 22 and 42. For ages 8, 9 . . . 13 the *Notes* add scores of 16, 18, 17, 15, 11, and 5, respectively, to the scores given by Ballard.

Burt: Northumberland Standardized Test III. Ages  $9\frac{1}{2}$ - $16\frac{1}{2}$ .

(University of London Press.)

Cattell: Intelligence Tests Scale II. Ages  $7\frac{1}{2}$ -adult.  
 Forms A and B.

Intelligence Tests Scale III. Ages  $12\frac{1}{2}$ -superior  
 adult. Forms A and B.

(Harrap.)

Dale: Bristol Group Reasoning Tests. Ages  $10\frac{1}{2}$ -14.  
 Forms A, B, and C.

(University of London Press.)

Fleming: Kelvin Measurement of Mental Ability. Ages 8-12.  
(Gibson.)

Group Test 33. Ages 14-superior adult.  
Percentile rankings for superior intelligence.  
(National Institute of Industrial Psychology.)

Group Test 34. Ages  $10\frac{1}{2}$ - $16\frac{1}{2}$ .  
Percentile rankings for three types of school.  
(National Institute of Industrial Psychology.)

Richardson: Simplex Junior Intelligence Scale. Ages 7-14.  
(Harrap.)

Schonell, F. J., and Adams, R. H.: The Essential Intelligence Test. Ages 7-14.  
(Oliver and Boyd.)

Spearman: A Measure of Intelligence for use in Schools. Ages 9-15.  
(Methuen.)

(c) GROUP TESTS. NON-VERBAL.

Alexander: Junior School Grading Test. Ages 5-12.  
Contains some verbal items.  
(University of London Press.)

Cattell: Intelligence Tests. Scale I. Ages 7-11. Forms A and B.  
(Harrap.)

Clark: Ryburn Group Intelligence Test. Ages  $7\frac{1}{2}$ - $11\frac{1}{2}$ .  
Partly verbal.  
(Gibson.)

Fleming: Kelvin Measurement of Ability in Infant Classes. Ages 5-8.  
(Gibson.)



N.I.I.P: Group Test 70/1. Ages 14 and over.

Group Test 70/23. Ages 14 and over.

(National Institute of Industrial Psychology.)

Mellone: Moray House Picture Intelligence Test I.

Ages  $6\frac{1}{2}$ – $8\frac{1}{2}$ .

(University of London Press.)

Otis: Group Intelligence Scale. Ages 6–9. Primary  
Forms A and B.

(Harrap.)

Sleight: Non-Verbal Intelligence Test. Ages 6–10.

(Harrap.)

(d) 11+ TESTS.

Some of the most satisfactory and best standardized intelligence tests are those devised for children of 11+. These tests are mainly of interest to Local Education Authorities. It is a common practice to use such tests in Transfer or Common Entrance Examinations, and the results are made available to the secondary schools. They are usually least discriminative for the dullest children, who would in general be better examined with non-verbal tests. For the sake of completeness the following list of 11+ tests is added. Probably this type of test is unsuitable for children with reading ages below 9 years.

Dawson: Mental Tests. Forms A and B.

(Harrap.)

Fleming and Jenkins: Cotswold Measurement of Ability.  
Series I and II.  
Series I in Welsh.

(Gibson.)

Houghton: Orton Intelligence Test.

(Gibson.)

Stephenson: Southend Group Test of Intelligence.

(Harrap.)

Thomson: Mental Survey Test 1932.

Moray House Tests.

(University of London Press.)

Northumberland Mental Test No. I and No. II.  
(Harrap.)

Tomlinson: Northern Test of Educability.

"West Riding" Tests of Mental Ability.  
Set Y and Set Z.

"West Yorkshire" Group Tests of Intelligence. Form X. Form W.  
(University of London Press.)

### III Special Mental Abilities Tests

Earle: Fife Test of Ability.	I. English (language).
Fife Test of Ability.	II. Science.
Fife Test of Ability.	III. Algebra.
Fife Test of Ability.	IV. Geometry.

Ages 11-12 $\frac{3}{4}$ .

Not attainment tests; designed to "differentiate between those pupils who seem to possess the ability in sufficient degree for successful further work in related subjects, and those pupils who seem to lack the ability in sufficient degree". To be given in first term in secondary school. Norms also include "probabilities" of pass or fail in both English and Scottish School Certificates for different score groups.

(University of London Press.)

Duplex Ability Test No. 1. Suitable for Age 10.  
Duplex Ability Test No. 2. Suitable for Age 11.  
Duplex Ability Test No. 3. Suitable for Age 12.  
Duplex Ability Test No. 4. Suitable for Age 13.



These tests give a measure of a child's "general intelligence" derived from the total score, while the scores on the sub-tests give useful indications of any differences between the efficiency of one type of mental process and another (e.g. verbal, logical, spatial).

The intelligence quotient is given and also its equivalent percentile rank. The percentile ranks of sub-tests are given. By allowing for different levels of general ability the test facilitates the selection of courses (linguistic, scientific, technical, commercial, craft, etc.).

The provision of separate tests for each age-group gives finer grading than would be possible with a test covering a wide age range. It also enables the author to dispense with timing, so that the slow but accurate worker gets full credit.

These tests appear to have great possibilities for Grammar and Technical Schools.

(Harrap.)

N.I.I.P: Form Relations Test (Space Perception).

Group Test 80 (Space Perception).

(National Institute of Industrial Psychology.)

Prak: Mathematical and Technical Tests. Ages 11-15.  
Tests A and B.

Standardized in Holland—English version. Non-verbal. Used for educational and vocational guidance.

(Harrap.)

Wing: Tests of Musical Ability and Appreciation.

Twenty-six tests of "musical intelligence" devised with the aim of providing a scientific set of psychological tests acceptable to psychologists, musicians, and teachers of music.

(Cambridge University Press.)

#### IV. Sensory-Motor Tests

The following are tests discussed in Chapter II.

Burt: Tests of Right and Left Handedness.

Tests of Eye Dominance.

*The Backward Child.*

#### V. Tests of Auditory Acuity

Burt: i. Watch Test.

ii. Speech Tests. (a) Vocalized. (b) Whispered.  
*The Backward Child.*

The above volume also discusses audiometers,  
both gramophone and pure-tone types.

Sheridan: Voice Test.

*The Child's Hearing for Speech.*

#### VI. Tests of Visual Acuity

Jaeger's Test Cards—for near vision.

Pray's Striped Letter Tests—for astigmatism.

Snellen's Letter Test—for distant vision.

Burt's *The Backward Child.*

#### VII. Tests of Colour Blindness

Collins and Drever: Group Tests for Colour Blindness—  
Publication IV of the Scottish Council  
for Research in Education.  
(University of London Press.)

### VIII. Interest and Personality Tests

Burt: Questionnaire on Neurotic Symptoms.

215 questions, many of which would be useful to teachers of adolescents.

*The Subnormal Mind.*

Cattell: Interest Test A.15. Age 14.

Divided into fifteen sections: travel; sports; commercial; mechanical; scientific; things of the mind; rural, naturalistic; religious; literary; artistic; decorative; sensual pleasure; sex; social; home. Separate norms for boys and girls and separate norms for each section which can therefore be given separately.

"F" Test.\*

Five tests: Pictures, Word Series, Completing Forms, Topics, Ink Blots. A temperament test, norms 11 to 14 years, but useful also for younger children from a diagnostic point of view.

Word Association Test.

100 association words especially arranged for children. Cattell also gives criteria of significant associations and Jung's original test.

*A Guide to Mental Testing.*

Hamley: Rating Scales.

For assessment of various qualities in connection with record cards.

A detailed scheme of educational guidance in a book full of insight into child nature.

*The Educational Guidance of the School Child.*



Lowenfeld: Mosaic Test.

For all ages. Gives indications of emotional disturbance.

(Institute of Child Psychology.)

Kaleidoblocs Test.

Not a "standardized" test, but suggestive for exploring character traits and mental abilities in a qualitative fashion.

(Badger Manufacturing Company,  
6 Pembridge Villas, London, W.11.)

Schonell, F. J.: Interest Schedule D.

Eighteen useful questions on the child's interests, with suggestions for the use of the resulting material with backward children.

Schedule B. Rating Scale for Temperament.

*Backwardness in the Basic Subjects.*

Valentine: Mental Types as revealed by Associations and Descriptions: Experiments XI, XII, and XIII.  
*An Introduction to Experimental Psychology.*

## BIBLIOGRAPHY

While the following list is in no way exhaustive, it contains all the volumes mentioned in footnotes throughout this book. Since an attempt has been made to interest teachers in all types of schools it is difficult to pick out a list of books equally useful to all. Those marked \* will probably be found useful in any school-teachers' library. Those marked \*\* will help the teacher who has an interest in statistics.

Alexander, W. P.: *The Educational Needs of Democracy*.  
(University of London Press.)

*Intelligence Concrete and Abstract*. (British Journal  
of Psychology; Monograph Supplement XIX.  
Cambridge University Press.)

Allport, G. W.: *Personality*. (Constable.)

Ballard, P. B.: *Group Tests of Intelligence*.\* (University  
of London Press.)

*The New Examiner*.\* (University of London Press.)

*Mental Tests*.\* (University of London Press.)

Boyd, W.: *The New Standard Spelling List*. (Harrap.)

Bühler, Charlotte: *From Birth to Maturity*. (Kegan Paul.)

Burt, Sir Cyril: *Mental and Scholastic Tests*.\* (Staples  
Press.)

*The Backward Child*.\* (University of London Press.)

*The Subnormal Mind*. (Oxford University Press.)

Cattell, R. B.: *A Guide to Mental Testing*.\* (University  
of London Press.)

Chaffe, G. H. (Ed.): *Careers Encyclopædia*. (Avon Press.)

- Drever, J., and Collins, M.: *Performance Tests of Intelligence*.  
(Oliver and Boyd.)
- Earle, F. M.: *Reconstruction in the Secondary School*.  
(University of London Press.)
- Fleming, C. M.: *Adolescence: Its Social Psychology*.  
(Routledge. Kegan Paul.)  
*The Beacon Arithmetic Teachers' Manual*. (Ginn.)
- Gardner, D. E. M.: *Testing Results in the Infant School*.  
(Methuen.)
- Gaw, F.: *Performance Tests of Intelligence*. (Industrial  
Fatigue Research Board, Report No. 31. H.M.  
Stationery Office.)
- Goodenough, F. L.: *Measurement of Intelligence by  
Drawings*. (World Book Club, Chicago.)
- Hamley, H. T. (Ed.): *The Testing of Intelligence*. (Evans  
Brothers.)
- Hamley, H. T. (and others): *The Educational Guidance of  
the School Child*. (Evans Brothers.)
- Hill, M. E.: *The Education of Backward Children*.  
(Harrap.)
- Hunt, E. P. A., and Smith, P.: *Scientific Vocational Guidance  
and its value to the choice of employment work of a  
L.E.A.* (City of Birmingham Education Com-  
mittee.)
- Inglis, A.: *Inglis Intelligence Quotient Values*. (Harrap.)
- Jacoby, H. J.: *The Analysis of Handwriting*. (Allen and  
Unwin.)
- Jung, C. G.: *Psychological Types*. (Kegan Paul.)
- Knight, R.: *Intelligence and Intelligence Tests*.  
(Methuen.)  
(G 906)



- Ministry of Education: *Pamphlet No. 5—Special Educational Treatment*. (H.M. Stationery Office.)
- Oakley, C. A., and Macrae, A.: *Handbook of Vocational Guidance*. (University of London Press.)
- Saudek, R.: *Experiments with Handwriting*. (Allen and Unwin.)
- The Psychology of Handwriting*. (Allen and Unwin.)
- Schonell, F. E., and Schonell, F. J.: *Diagnostic and Attainment Testing*. (Oliver and Boyd.)
- Schonell, F. J.: *The Diagnosis of Individual Difficulties in Arithmetic*.\* (Oliver and Boyd.)
- The Psychology and Teaching of Reading*.\* (Oliver and Boyd.)
- Backwardness in the Basic Subjects*.\* (Oliver and Boyd.)
- Essentials in Teaching and Testing Spelling*.\* (Macmillan.)
- Scottish Council for Research in Education: *Studies in Reading, Vol. I*. (University of London Press.)
- Sheridan, M. D.: *The Child's Hearing for Speech*. (Methuen.)
- Terman, L. M., and Merrill, M. A.: *Measuring Intelligence*. (Harrap.)
- Thomas, T.: *The Science of Marking*\*\* (John Murray.)
- Valentine, C. W.: *Intelligence Tests for Children*.\* (Methuen.)
- An Introduction to Experimental Psychology*. (University Tutorial Press, London.)
- Vernon, P. E.: *The Measurement of Abilities*\*\* (University of London Press.)
- The Standardization of a Graded Word Reading Test*.\* (University of London Press.)

- Watts, A. F.: *The Language and Mental Development of Children*.\* (Harrap.)
- Williams, G. Perrie: *The Northamptonshire Composition Scale*. (Harrap.)
- Wing, H.: *Tests of Musical Ability and Appreciation*. (British Journal of Psychology; Monograph Supplement No. XXVII. Cambridge University Press.)

## APPENDIX I

### THE EDUCATIONALLY SUBNORMAL—SPECIAL EDUCATIONAL TREATMENT

The term "special educational treatment" is one which acquired a great deal of significance with the passing of the Education Act, 1944. This is the term used in the Act with reference to the sort of education needed by those whom the Ministry of Education in its Regulations has called "Handicapped Pupils". These include the blind, the partially sighted, the deaf, the partially deaf, the delicate, the diabetic, the educationally subnormal, the epileptic, the maladjusted, the physically handicapped, and those with speech defects. These handicaps are defined in the Ministry's "Pamphlet No. 5—Special Educational Treatment", which also discusses "ascertainment", estimated proportions of the various categories to be found in the school population, and how and where the special educational treatment should be given.

The educationally subnormal are defined as those who "by reason of limited ability or other conditions resulting in educational retardation, require some specialized form of education wholly or partly in substitution for the education normally given in ordinary schools". A fair proportion of the pamphlet consists of a discussion of this problem, and all teachers, in primary schools at least, should make themselves conversant with it. The definition of educational subnormality is not at all specific, but the Ministry's suggestion is that children should not come into this category unless they are retarded in their school work by more than 20% of their age, that is unless they have an educational quotient below 80.

On the other hand, they must not be of such low grade as



to be ineducable or have a distracting influence on others. It is suggested that children whose intelligence quotient is below about 55 cannot be educated at school. Such children should be reported to the Local Authority, to be dealt with under the Mental Deficiency Acts. It should be noted that they, if certified, are now the only mentally defective children in Great Britain, for the term M.D. is no longer applicable to any *school* child. In any case it has always been a legal and not a psychological or educational term. The Ministry deprecates the practice of retaining, on sentimental grounds, children who are ineducable or whose presence in school is detrimental to others. The Ministry points out the distraction such children can cause and the attention which they require from the teacher, and it makes its intentions quite clear in a statement which would have seemed like a charter of emancipation to many a teacher in the past, namely, "In future this practice should cease".

### Placement of the Educationally Subnormal.

Educable children of low intelligence, that is, the educationally subnormal, are generally children with intelligence quotients of from 55-75, and should be provided for by one of three types of organization, Boarding Special Schools, Day Special Schools, or by special educational treatment in ordinary schools.

*Boarding Special Schools* are needed for those educationally subnormal children who for their own good require to be taken away from home. Among these will be found chronic truants, unmanageable children, children from unsuitable homes, and those committed to the care of the Local Authority by the Courts. Approximately two children out of every thousand are in this category.

*Day Special Schools* provide special educational treatment for the more seriously dull who do not require boarding school treatment, usually all those with intelligence quotients of from 55 to 65, together with *certain* children with intelligence

quotients running up to 75. These will be the unstable, the irregular attenders, and others who are so discouraged that they are obviously not profiting, and are not likely to profit, by education in the ordinary school. Because of its variability (see Chapter II) the intelligence quotient should not be the sole criterion in the selection of children for these special schools. Even a child with an intelligence quotient as low as 65 may have a stable temperament and other favourable factors in his make-up that will enable him to fit quite well into the ordinary school, that is, assuming the school can provide suitable education for him. The Ministry expects to find about 1% of children requiring the Day Special School service, but teachers in schools known to contain a high proportion of dull children should remember (vide Chapter VI) that this percentage may be much higher in their particular schools.

*Ordinary Schools* are expected to deal with the remainder of the educationally subnormal, and since these are estimated to be 10% of the total school population, the average Primary School will have about 9% where there is adequate Day Special School provision. Individual schools will have more, or less, than this proportion, and most Secondary Modern Schools will have more. The Ministry is in no way dogmatic about the type of educational treatment to be arranged for these children, but it will clearly have a close relation to school organization. Whether they are dealt with in "special classes", in the lower "stream", or in some other way, will depend on the size of the school, the size of the staff, and the available accommodation.

### Ascertainment.

This clearly involves the co-operation of a number of people, such as teachers, school medical officer, psychologist, and perhaps the children's officer, probation officer, or other person competent to give advice. For the educationally sub-normal child who is capable of being retained in the ordinary



school this will usually be the responsibility of the teachers, aided possibly by an educational psychologist or other officer appointed for the purpose by the Local Education Authority. Where a Special School is concerned, however, the Local Authority has the final decision, being advised by an approved school medical officer after he has examined the child and been advised by teacher, psychologist, or other specialist. There are safeguards of the parents' interests, and an appeal can be made to the Minister. On the other hand, if everyone from the Minister downwards is satisfied that it is in the child's best interest to attend a Special School the law can be invoked to see that the parent allows the child to attend.

Any teacher who has good grounds for believing that a child needs special educational treatment in a Special School should at once report the matter to the appropriate authority. The tests described in this book will guide the teacher in coming to this conclusion. The Local Authority will then require the head teacher of the school to complete a form<sup>1</sup> prescribed by the Ministry's regulations, as a preliminary to the child's examination by an approved medical officer.

This report form asks for certain routine information, including particulars of the child's school attendance, and for information about his scholastic attainments, behaviour and disposition, and home environment. Readers of the foregoing chapters who have to complete this form may be interested to find that it requires them to administer two tests, which can be recognized as the earlier portions of Burt's word reading and dictation tests, and that they are also asked to append the child's drawing of a man.

Not all Local Authorities have yet been able to make the Special School provision required of them by the 1944 Act, and there are admitted difficulties in rural areas. Nevertheless, it is the teacher's duty, in my opinion, to call the Authority's attention to any children for whom he considers Special School provision should be made. He should also report the

<sup>1</sup> Form 3 H.P.



lowest grade children with a view to possible exclusion from school.

It is not usual to make any decision about educational subnormality until the child reaches the age of 7 years, unless mental deficiency is suspected, but the Authority should be prepared to examine a child before this age in order to guide the teacher on the possible later decision. There is, however, more than sufficient test material suggested in this book to help the teacher to form a rough conclusion without resorting to the Authority.

## APPENDIX II

### VOCATIONAL GUIDANCE

In spite of some excellent pioneering in various parts of the country the vocational guidance service, from the psychologist's point of view, is still in an embryo stage so far as Great Britain as a whole is concerned. Nevertheless, the last twenty-five years have seen a growing recognition of the valuable part which psychological tests can play in giving good vocational guidance.

This subject has two important aspects: (a) the mental and physical qualities possessed by the young people needing guidance, and (b) the particular abilities and qualities, physical and temperamental, required for success in the employments available.

Some schools have appointed "careers masters". While one should be careful that this work is not undertaken by people with more enthusiasm than knowledge, there is a definite place in it for teachers who are interested in tests and who are prepared to study the subject of vocational guidance. In a published report on a research into the value of scientific vocational guidance conducted by a Local Education Authority<sup>1</sup> a recommendation is made that training courses should be arranged for selected teachers so that each secondary school would have at least one person competent to administer tests and collect relevant data. It is further suggested that these teachers should work in co-operation with an officer qualified to select tests and advise on methods, so that

<sup>1</sup> *Scientific Vocational Guidance and its value to the choice of employment work of a L.E.A.*

cumulative record cards could be compiled to be used at the end of each child's school life by the juvenile employment officer in co-operation with the head and the trained teacher. This report, incidentally, shows that, after a very adequate follow-up, the procedures adopted, which included a comprehensive testing programme, were indeed giving valuable guidance, judged from several different points of view.

Four batteries of tests were used to measure respectively performance ability, manual dexterity, mechanical ability, and dressmaking ability. An intelligence test and a test of clerical ability were also used. These tests were given in conjunction with rating scales for numerous subjects. There were also a special medical report, a report on the home, and an interview. The complete report on the research is well worth studying.

Consideration of vocational guidance soon brings to light the not entirely negative virtue it possesses of steering a child away from an unsuitable job. In this connection the intelligence test is a useful "sieve" for preventing the more obvious misfits. Cattell has compiled a list of "Occupational Norms of Intelligence"<sup>1</sup> which gives both the average intelligence quotient and the scatter of intelligence over the middle 50% of people in each occupation listed. Any youngster proposing to enter one of these occupations with insufficient intelligence should generally be advised against such a choice, unless there are other known factors pointing in a different direction. The question of the possession of too much intelligence for a particular job must also be considered.

The ability to estimate what level of intelligence and attainment and what special abilities and temperamental qualities are needed *for the job in question* is, of course, a pre-requisite for vocational guidance. A great deal of research has still to be done, though the list of occupational requirements com-

<sup>1</sup> *A Guide to Mental Testing*. These norms were drawn up by means of Cattell's Intelligence Scale III, and this test should be used for strict comparison.



piled by Oakley and Macrae<sup>1</sup> provides a solid contribution in this direction. These authors have listed upwards of a hundred occupations, and for each have indicated minimal abilities and qualities necessary under twenty-three different heads. They discuss intelligence and special abilities ranging from "school level" to "attractiveness of personal appearance", and give an account of test material. The book contains notes on training requirements for the various occupations, the names of appropriate professional and other bodies, and the numbers entering each vocation annually. (These last will be somewhat out-of-date now, at least in the 1937 edition of their book. It will also be found that the information relates in the main to children of above average intelligence.)

A more recent publication is the *Careers Encyclopædia*, edited by G. H. Chaffe. This work has not the scientific basis of the work of Oakley and Macrae, but it is more up-to-date. It deals with 220 occupations, describing the nature of these, and discussing age of entry, educational requirements, and training necessary. It also deals with openings, salaries, and prospects in the various occupations studied.

Booklets on vocational matters may be obtained from the Ministry of Labour. Some Local Education Authorities have also published excellent notes about jobs available in their areas. Some teachers may live in areas where the choice of employment is not wide. In such cases it should not be impossible, given the co-operation of employers, to compile useful information concerning occupational requirements. These areas, where the population is largely absorbed by one or two industries, provide an excellent field for small pieces of research, in which the volumes mentioned in this appendix would be useful guides.

The modern practice of arranging visits by children to local industries is excellent. It would be a very good thing if all secondary school teachers could be included in these visits at one time or another, so that they could see the types

<sup>1</sup> *Handbook of Vocational Guidance.*

of work their pupils will be doing in the future. Tests such as those referred to in the volumes on vocational guidance already mentioned, scholastic tests, various special tests dealt with in the body of this book, and others, perhaps yet to be devised, should all be of considerable value in the task of placing young people in jobs where they will be both useful and happy. These tests, however, are not sufficient in themselves, for vocational guidance is as much an art as a science, and an understanding of life and the work of man will always be an essential factor in its application.

## INDEX

- ABILITY, 10, 11, 54.
  - distribution of, 104, 107 ff.
  - tests, 20.
- Achievement, 10, 11.
- Age factor, 25, 32.
- Age-group records, 137, 138.
  - spread, 139.
- Algebra, standardized test of, 154.
  - Test of Ability for, 159.
- Arithmetic, composite, 153.
  - difficulties in, 88.
  - mental, 151.
  - mechanical skill in, 151, 152.
  - problem, reasoning, or method, 152.
  - tests, 15, 59, 60, 75, 76.
  - "weakness" figure in, 97.
- Attainment ages, 10.
  - norms, 145.
  - range, 69.
  - tests, 10, 31, 58, 68, 147.
- Attainments, distribution of, 126.
- Auditory acuity, 161.
- Average intelligence, 107, 108.
  
- BASIC NUMBER COMBINATIONS, 95.
- Beacon Arithmetic Diagnostic Tests, 94.
  
- CASE RECORDS, 101.
- Categories, value of, 6.
- Character assessments, 67.
  - tests, 65.
  - traits, 127.
- Chemistry tests, 154.
- Class surveys, 14 ff.
  
- Class records, 131, 132, 134 ff.
- Coaching, 37.
- Colour blindness, 64.
  - tests, 161.
- Common Entrance Examination, 3, 4, 23, 37, 69, 114.
- Common sense, 33.
- Correlation, 32.
- Courses, selection of, 20, 159, 160.
- Craft work, 18.
  
- DEAFNESS, HIGH-FREQUENCY, 63, 93, 101.
- Diagnosis, 20, 21.
- Diagnostic tests, 88.
  - Beacon Infant Reader, 101.
  - of arithmetic, 93, 94.
  - of reading, 99.
  - of spelling, 98.
  - of written English, 101 ff.
- Dictation tests, 98, 151.
- Distribution, normal curve of, 110, 111.
- Drawing test, 155.
- Dull children, 104.
  
- EDUCATIONAL RESEARCH, 17, 18.
  - history, 101.
- Educationally subnormal, 2, 8, 13, 168 ff.
  - ascertainment of, 170, 171.
  - definition of, 168.
  - placement of, 169.
- Eleven plus tests, 158.
- English, Test of Ability for, 159.
  - tests, other than reading, 149, 150.



- English, written, 101 ff.  
 Examiner's instructions, 78.  
 Eye dominance, 62, 91.  
   tests, 161.
- "F." (FLUENCY OF ASSOCIATION) TEST, 66.  
 Formboard test, 156.
- GEOGRAPHY TESTS, 154.  
 Geometry test, 154.  
   Test of Ability for, 159.  
 Grammar, 18.  
 Grammar Schools, Secondary,  
   *see* Secondary Schools.  
 Graphology, 61, 67.  
 Group tests, 48, 49, 156, 157.
- "HALO EFFECT", 31.  
 Handedness, 62, 91.  
   test, 161.  
 Handwriting, 60, 61, 67.  
   tests, 153.  
 "Headroom", 29.  
 Hearing tests, 161.  
 Histogram, 111.  
 History tests, 161.
- IMMEDIATE RECALL, TESTS OF, 98.  
 Individual records, 130, 137.  
   tests, 46, 47, 80, 81, 154 ff.  
 Ineducable child, 169.  
 Inhibited child, 66.  
 Intellect, 104.  
 Intelligence and social level, 121.  
   definition of, 32, 42.  
   distribution of, 119, 120, 124, 125.  
   growth of, 41, 107.  
   tests, 10, 31, 32, 46, 55, 68, 77, 84, 105, 154 ff.  
   variability of, *see* Intelligence quotient, constancy of.  
 Intelligence quotient, 2, 32.  
   constancy of, 38 ff.  
   limitations of, 2.
- Intelligence quotients, distribution of, 42.  
   mean, 122, 123.  
   proportion of, 111, 112.  
 Interest tests, 65.  
   and personality tests, 162, 163.  
 Intuition, 92.
- JUDGMENTS, SUBJECTIVE, 23.
- LANGUAGE TESTS, 73  
 "Late bloomer", 106.
- MATHEMATICAL TESTS, 54, 160.  
 Matrices test, 155.  
 Maturation level, 9, 17 (*see also* Reading, readiness).  
 Maze test, 155.  
 Median, 132, 133, 138, 139.  
   mark, 133.  
 Mental age, 32, 68 ff.  
   deficiency, 104.  
   ratio, 38.  
   types, 66.  
 Modern Schools, Secondary, *see* Secondary Schools.  
 Musical ability, 57.  
   tests, 160.
- NON-VERBAL TESTS, 50, 51, 53, 79, 155, 157.  
 Norms, 16.
- OCCUPATIONAL NORMS OF INTELLIGENCE, 174.  
   requirements, 174, 175.
- PERCENTILE RANKS, 45, 112, 117, 118.  
 Performance tests, 54 ff., 154, 155.  
 Personality, assessment of, 67.  
 Phonic units, analysis of, 99.  
 Policy making, 14 ff.  
 Prognosis, 19, 20.

QUARTILE GROUPS, 126.  
marks, 113, 123.  
Quartiles, 132, 133, 138, 139.

RATING SCALES, 66, 67, 126, 127,  
162, 163.

Reading, 72, 99 ff.  
readiness, 6, 105.  
tests, 15, 28, 58, 71 ff., 100,  
147, 148.

Reasoning tests, 155, 156.

Record cards, 67.

Reliability, 28, 36, 44.  
coefficient, 101.  
of teachers' assessments, 66.

SCHOLARSHIP EXAMINATION, *see*  
Common Entrance Exami-  
nation.

Schonell Diagnostic Arithmetic  
Tests, 94.

School marks, 113.  
surveys, 14 ff., 142 ff.

Science, Tests of Ability for, 159.

Script Writing Scale, 153.

Secondary Schools, 18, 20, 54,  
70, 81 ff., 82.

Grammar, 3 ff., 81 ff., 111,  
112.

Modern, 18, 70, 83 ff., 91,  
136 ff.

Technical, 5, 54, 83.

Sensory tests, 62, 101.

Sensory-motor tests, 62, 63,  
161.

Seguin Formboard, 54, 45.

Space perception tests, 160.

Special abilities, 54.

Educational Treatment, 168.  
mental abilities tests, 159, 160.

Schools, 169, 170.

Specific abilities, 33.

disabilities, 69, 91, 92.

Spelling, 97.

Spelling schemes, 74, 75.  
tests, 24 ff., 58, 59, 74, 98, 150,  
152.

Standard deviation, 43, 44, 114 ff.

Standardized tests, diagnostic  
possibilities of, 89, 90.  
use of, 27, 28.

TEACHERS' ASSESSMENTS, 66, 67.  
difficulties, 12.

Technical Schools, *see* Sec-  
ondary Schools.

tests, 54, 160.

Temperament, relating scales for,  
67.

Terman-Merrill test, 33.

Test data, 69, 128, 146.

Test results, use of, 129.  
reliability of, 28, 36, 44.  
validity of, 28, 35, 36.  
value of, 128.

Tests, administration of, 52, 53.  
as administrative devices, 5, 9.  
choice of, 68 ff.  
criticisms of, 26, 29, 34, 35.  
effective use of, 72.  
standardization of, 24 ff.

Timed tests, 29.

Transfer Examination, *see* Com-  
mon Entrance Examination.

VALIDITY COEFFICIENT, 101.

Verbal tests, 50, 51, 53, 79.  
defects of, 51, 70, 73, 74.

Vision, defective, 64.

Visual acuity tests, 161.

Vocabulary tests, 26, 84, 90,  
149.

Vocational guidance, 173 ff.

WORD DISCRIMINATION TEST,  
100.

Words, directional attack on, 99.

Form No. 3.

PSY, RES.L-1

**Bureau of Educational & Psychological  
Research Library.**

The book is to be returned within  
the date stamped last.

.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....
.....	.....	.....



1

---

# THE TECHNIQUE OF TEACHING

A new series of books on subjects of vital  
interest to all teachers

1. THE PRINCIPLES  
OF ARITHMETIC

By Alexander Keith, M.A., B.Sc., and  
James Robertson, M.A., D.Sc.

2. GAMES ACTIVITIES  
FOR GIRLS

By Margaret Dunn

3. FROM DAY TO DAY IN  
THE INFANT SCHOOL

By F. Irene Serjeant, N.F.U.T.D.

4. THE APPLICATION OF  
PSYCHOLOGICAL TESTS  
IN SCHOOLS

By Noel E. Whilde, B.Sc.

5. SPEECH IN SCHOOLS

By F. J. Griffiths

6. THE TEACHING OF ART

By L. de C.-Bucher, A.R.C.A.

7. PLAYGROUND GAMES FOR  
SECONDARY BOYS

By L. C. Williams and  
A. W. Willee, B.A.

